

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2003, 08:18:04 ; Search time 17 Seconds
(without alignments)
365.149 Million cell updates/sec

Title: US-09-877-160A-1

Perfect score: 662

Sequence: 1 MKFTIAFFVATLAVMTVSGE.....GDILERDLKKBARKVKIEV 132

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 segs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	657	99.2	132	1	P14004 dermatophag
2	273.5	41.3	134	1	O96870 blomia trop
3	187	28.2	110	1	Q94522 lepidoglyph
4	90	13.6	886	1	O33600 sulfolobus
5	88	13.3	585	1	P38748 saccharomyc
6	87.5	13.2	2230	1	Q13439 homo sapien
7	87	13.1	757	1	Q8W3W3 homo sapien
8	86.5	13.1	978	1	O67124 aquifex aeo
9	85.5	12.9	254	1	Q9V7Z6 drosophila
10	85	12.8	784	1	P08177 escherichia
11	85	12.8	1433	1	O42184 gallus gall
12	84	12.7	166	1	P75260 mycoplasma
13	83.5	12.6	603	1	O51413 borrelia bu
14	83	12.5	2245	1	P54697 dictyostell
15	82.5	12.5	878	1	Q8RAH4 thermoaer
16	82.5	12.5	1978	1	P10587 gallus gall
17	82	12.4	144	1	O30189 archaeoglob
18	81.5	12.3	882	1	P58301 pyrococcus
19	81	12.2	886	1	O22320 archaeoglob
20	80.5	12.2	2442	1	Q9BV73 homo sapien
21	80	12.1	691	1	P39814 bacillus su
22	80	12.1	784	1	P46067 erwinia amy
23	80	12.1	1972	1	P35749 homo sapien
24	79.5	12.0	565	1	Q9L974 bacillus ha
25	79.5	12.0	1875	1	Q02455 saccharomyc
26	79.5	12.0	2869	1	Q00798 plasmodium
27	79	11.9	559	1	Q43207 triticum ae
28	78.5	11.9	436	1	Q00387 xenopus lae
29	78.5	11.9	578	1	P46150 drosophila
30	78.5	11.9	705	1	P34562 caenorhabdi
31	78.5	11.9	845	1	G60563 mesocricetu
32	78	11.8	879	1	Q96YR5 sulfolobus
33	78	11.8	914	1	Q9UY36 pyrococcus

34	78	11.8	1972	1	MYHB_PABIT	P35748 oryctolagus
35	77.5	11.7	1055	1	TIRH_HAEIN	O05052 haemophilus
36	77	11.6	284	1	TPMM_ANISI	Q9NAS5 anisakis si
37	77	11.6	284	1	TPMM_TRICO	P15846 trichostrom
38	77	11.6	2017	1	MYSN_DROME	O99323 drosophila
39	76.5	11.6	225	1	YG54_METJA	O59048 methanococ
40	76.5	11.6	419	1	NEMO_HUMAN	Q9Y6K9 h nf-kappab
41	76.5	11.6	419	1	MEP1_ARATH	Q9LW85 arabidopsis
42	76.5	11.6	984	1	SECA_AQUAE	O67718 aquifex aeo
43	76.5	11.6	1072	1	CARB_THETN	Q8Rbk0 thermoanaer
44	76.5	11.6	1976	1	MYHA_BOVIN	Q27991 bos taurus
45	76.5	11.6	1976	1	MYHA_HUMAN	P35580 homo sapien

ALIGNMENTS

RESULT 1

ALL5_DERPT
ID ALL5_DERPT STANDARD; PRT; 132 AA.
AC P14004;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mite allergen Der p 5 (Der p V) (IgE-binding allergen).
GN DERP5.
OS Dermatophagoides pteronyssinus (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcotiformes; Astigmata; Psoroptidia; Analgoidea;
OC Pyroglyphidae; Dermatophagoides.
OX NCBI_TaxID=6956;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95096428; PubMed=7798547;
RA Lin K.B., Hsieh K.H., Thomas W.R., Chiang B.L., Chua K.Y.;
RT "Characterization of Der p V allergen, cDNA analysis, and IgE-mediated
reactivity to the recombinant protein.";
RL J. Allergy Clin. Immunol. 94:989-996(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90010801; PubMed=2794865;
RA Tovey E.R., Johnson M.C., Roche A.L., Cobon G.S., Baldo B.A.;
RT "Cloning and sequencing of a cDNA expressing a recombinant house dust
mite protein that binds human IgE and corresponds to an important low
molecular weight allergen.";
RL J. Exp. Med. 170:1457-1462(1989).
RN [3]
RP REVISION TO 132.
RA Cobon G.S.;
RL Submitted (FEB-1990) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE MITE GROUP 5 ALLERGEN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; S76337; AAB32841.1; --
DR EMBL; S76340; AAB32842.1; --
DR EMBL; X17699; CAA35692.1; ALT_INIT.
KW Allergen.
FT VARIANT 61 61 E -> A.
SQ SEQUENCE 132 AA; 15615 MW; 3C840C3C26707C05 CRC64;

Query Match 99.2%; Score 657; DB 1; Length 132;
Best Local Similarity 99.2%; Pred. No. 2.2e-43;
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MKFTIAFFVATLAVMTVSGEDKXHDYQNEFDLLMERIHQIKKGELALFYLOQINHEE 60
|||||

```

Db 1 MKFIIAFFVATLAVMTVSGDKKHQYQNEFDLLMERIHEQIKKGELALFYQEQINHF 60
QY 61 EKPTKEMDKIVAEMDTIIAMIDGVGVLDRLMQRKDLDFEQYNLEMLKSGDILERDL 120
Db 61 EKPTKEMDKIVAEMDTIIAMIDGVGVLDRLMQRKDLDFEQYNLEMLKSGDILERDL 120
QY 121 KKEEARVKKIEV 132
Db 121 KKEEARVKKIEV 132

RESULT 2
ALL5 BLOTA
ID ALL5 BLOTA STANDARD; PRT; 134 AA.
AC O96870; Q17283;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mite allergen Blo t 5.
GN BLOT5.
OS Blomia tropicalis (Mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcoptiformes; Astigmata; Glycyphagoidea;
OC Echinopodidae; Blomia.
OX NCBI_TaxID=40697;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=97154616; PubMed=9001334;
RX Artuda L.K., Vailles L.D., Platts-Mills T.A.E., Fernandez-Caldas E.,
RA Moncealegre F., Lin K.-L., Chua K.-Y., Rizzo M.C., Nasipitz C.K.,
RA Chapman M.D.;
RT "Sensitization to Blomia tropicalis in patients with asthma and
RT identification of allergen Blo t 5.";
RL Am. J. Respir. Crit. Care Med. 155:343-350 (1997).
RN [2]
RP SEQUENCE OF 63-134 FROM N.A.
RX MEDLINE=96426241; PubMed=8828535;
RA Caraballo L., Avjoglou A., Marrugo J., Puerta L., Marsh D.;
RT "Cloning and expression of complementary DNA coding for an allergen
RT with common antibody-binding specificities with three allergens of the
RT house dust mite Blomia tropicalis.";
RL J. Allergy Clin. Immunol. 98:573-579 (1996).
CC -!- SIMILARITY: BELONGS TO THE MITE GROUP 5 ALLERGEN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U59102; AAD10850.1; -
DR EMBL; U27702; AAB49396.1; -
KW Allergen.
FT CONFLICT 63 82 NKSKELOEKIIRLDVVCAM -> RAKELQRRKSPENLMWF
FT APL (IN REF. 2).
FT CONFLICT 126 126 Q -> P (IN REF. 2).
FT SEQUENCE 134 AA; 15642 MW; F26D763475C855FD CRC64;

Query Match 41.3%; Score 273.5; DB 1; Length 134;
Best Local Similarity 42.5%; Pred. No. 2.2e-14;
Matches 57; Conservative 32; Mismatches 40; Indels 5; Gaps 2;

QY 1 MKF---ITAFFVATLAVMTVSGDKKHQYQNEFDLLMERIHEQIKKGELALFYQEQIN 57
Db 1 MKFAVLVIACFAA--SVLAQEHKPKKDDFRNEFDLLIEQANHALEKGEHQLLYLQHQLD 58
QY 58 HFEEKPTKEMDKIVAEMDTIIAMIDGVGVLDRLMQRKDLDFEQYNLEMLKSGDILE 117
Db 59 ELNENKSELOEKIIRLDVVCAMIEGAGGALERELKRTDLNILDRFNRYEQAUTLSKILL 118
QY 118 RDLKKEEARVKKIE 131

```

```

Db 119 KDLKETQKVXDIQ 132
:||||:|:||||:|:

ALL5 LEPDS
ID ALL5 LEPDS STANDARD; PRT; 110 AA.
AC Q9USE2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mite allergen Lep d 5 (Fragment).
OS Lepidoglyphus destructor (Storage mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcoptiformes; Astigmata; Glycyphagoidea; Glycyphagidae;
OC Lepidoglyphus.
OX NCBI_TaxID=36936;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=2109292; PubMed=11168362;
RX Eriksson T.L.J., Rasool O., Ruecas S., Whitley P., Cramer R.,
RA Appenzeller U., Gafvelin G., van Hage-Hamsten M.;
RT "Cloning of three new allergens from the dust mite Lepidoglyphus
RT destructor using phase surface display technology.";
RL Eur. J. Biochem. 268:287-294 (2001).
CC -!- SIMILARITY: BELONGS TO THE MITE GROUP 5 ALLERGEN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ250278; CAB62212.1; -
KW Allergen.
FT NON TER 1
FT SEQUENCE 110 AA; 12550 MW; D405050089E24CF1 CRC64;

Query Match 28.2%; Score 187; DB 1; Length 110;
Best Local Similarity 36.4%; Pred. No. 6.2e-08;
Matches 39; Conservative 24; Mismatches 44; Indels 0; Gaps 0;

QY 25 DYQNEFFLLMERIHEQIKKGELALFYQEQINHFPEKPTKEMDKIVAEMDTIIAMIDG 84
Db 2 DFRNEFDRLIHTMEQFPAKLEQALHLSHQVTELEKSKELKAQILREISIGLDFIDS 61
QY 85 VRGVLDRLMQRKDLDFEQYNLEMLKSGDILERDLKKEEARVKKIE 131
Db 62 AKGHFERELKADLNLAKEKFNFSALSTGAVLHKDLTALATKVAIE 108

RESULT 4
ALL5 SULAC
ID RASO SULAC STANDARD; PRT; 886 AA.
AC O33600;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA double-strand break repair rad50 ATPase.
GN RAD50.
OS Sulfolobus acidocaldarius.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2285;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=ATCC 33909 / NCIB 11770 / DSM 639;
RX MEDLINE=97362314; PubMed=9211741;
RA Elie C., Baucher M.F., Fondrat C., Forterre P.;
RT "A protein related to eucaryal and bacterial DNA-motor proteins in the
RT hyperthermophilic archaeon Sulfolobus acidocaldarius.";

```

```

RL J. Mol. Evol. 45:107-114(1997).
CC -!- FUNCTION: Involved in DNA double-strand break repair (DSBR). The
CC rad50/mre11 complex possesses single-strand endonuclease activity
CC and ATP-dependent double-strand-specific exonuclease activity.
CC Rad50 provides an ATP-dependent control of mre11 by unwinding
CC and/or repositioning DNA ends into the mre11 active site (By
CC similarity).
CC -!- SUBUNIT: Forms a complex with mre11 (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y10687; CAA71688.1; -.
CC HAMAP; MF_00449; -. 1.
CC InterPro; IPR003439; ABC transporter.
CC InterPro; IPR003395; SMC N.
CC Pfam; PF04423; Rad50_zn_hook; 1.
CC Pfam; PF02463; SMC_N; 1.
CC ProDom; PD000006; ABC transporter; 1.
CC DNA repair; Hydrolyase; ATP-binding; Coiled coil.
CC NP BIND 30 37 ATP (BY SIMILARITY).
CC DOMAIN 174 727 COILED COIL (POTENTIAL).
CC SEQUENCE 886 AA; 103857 MW; 0390AE1403194104 CRC64;
CC
CC Query Match 13.6%; Score 90; DB 1; Length 886;
CC Best Local Similarity 21.6%; Pred. No. 11;
CC Matches 32; Conservative 37; Mismatches 43; Indels 36; Gaps 6;
CC
CC 20 EDKQHDYQNEFDLMERLHEQIKKGELALP--YLQE---QINHPERPTK----- 65
CC Db EDELQYNTENRIKEIKVQVDILSGELSVNKKIEIALRLKDFEKEKRYNKIETRVK 282
CC
CC 66 ---EMKDKI---VAEMDTIAMDVGVRGVL-----DLMQRKDLDIFQYNLEML 109
CC Db ELDENREKINTISSFKSLVQIDSLSKQINNVENDLKKRKEKRLKKELEKEKQYERIE 342
CC
CC 110 KSGDILDERDLKKEAR-----VKKIE 131
CC Db KRKKELEKEKQYERIEKRLITVLEKNIE 370
CC
CC RESULT 5
CC YHBO YEAST STANDARD; PRT; 585 AA.
CC AC
CC DT 01-FEB-1995 (Rel. 31, Created)
CC DT 01-FEB-1995 (Rel. 31, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Hypothetical 67.5 kDa protein in PRP94-STE20 intergenic region.
CC GN YH1010.
CC OS Saccharomyces cerevisiae (Baker's yeast).
CC OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
CC OX NCBI_TaxID=4932;
CC RN
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=S288c / AB972;
CC RX MEDLINE=94378003; PubMed=8091229;
CC RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
CC Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,
CC RA Kucaba T., Hallier L., Jier M., Johnston L., Langston Y.,
CC RA Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
CC RA Nhan M., Rifkin L., Riles L., St Peter H., Trevisan S., Vaughan K.,
CC RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
CC RA Vaudin M.;
CC RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
CC VIII.";
```

```

RL Science 265:2077-2082(1994).
CC -!- SIMILARITY: TO C.ELEGANS REED9.9 AND S.POMBE SPAC16E8.13.
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U11582; AAB65064.1; -.
CC PIR; S46825; S46825
CC SGD; S0001002; YHL010C.
CC InterPro; IPR001841; Znf_ring.
CC InterPro; IPR001607; Znf_UBP.
CC Pfam; PF00097; ZF-C3HC4; 1.
CC Pfam; PF02148; ZF-UBP; 1.
CC SMART; SM00184; RING; 1.
CC SMART; SM00290; Znf_UBP; 1.
CC PROSITE; PS00518; ZF_RING_1; FALSE_NEG.
CC PROSITE; PS00089; ZF_RING_2; 1.
CC KW Hypothetical protein; Zinc-finger.
CC FT ZN FING 240 280 RING-TYPE.
CC SQ SEQUENCE 585 AA; 67503 MW; 88FF670CC73A1263 CRC64;
CC
CC Query Match 13.3%; Score 88; DB 1; Length 585;
CC Best Local Similarity 23.7%; Pred. No. 11;
CC Matches 31; Conservative 23; Mismatches 47; Indels 30; Gaps 3;
CC
CC 18 SGEQKHQYQNEPFDLML-----ERIHEQIKKGELALFYLOEQINHPERPTKEMK 68
CC Db 376 SGDDNDNDIGNSDELQNVVYGNRSNKGESNKNKDGELAAFLRHRHYHLE----- 427
CC
CC 69 DKVAEMDTIAMDVGVRGVLDRMQRKDLDFQYNLEMLKSGDILDERL----- 120
CC Db 428 ----YQVLISQLESQREYVELKLEQKQQTASDSSNVESLKKSMDLKLQFQVTKEMW 482
CC
CC 121 KKEARVKKIE 131
CC Db 483 KREMAQSKLE 493
CC
CC RESULT 6
CC GOG4 HUMAN STANDARD; PRT; 2230 AA.
CC ID
CC AC Q13439; Q13270; Q13654; Q14436;
CC DT 16-OCT-2001 (Rel. 40, Created)
CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
CC DE Golgi autoantigen, golgin subfamily A 4 (trans-Golgi p230) (256 kDa
CC DE golgin) (Golgin-245) (72.1 protein).
CC GN GOLGA4.
CC OS Homo sapiens (Human).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC OX NCBI_TaxID=9606;
CC RN
CC RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
CC RX MEDLINE=96215236; PubMed=8626529;
CC RA Erlich R., Gleeson P.A., Campbell P., Dietzsch E., Toh B.-H.;
CC RT "Molecular characterization of trans-Golgi p230: a human peripheral
CC RT membrane protein encoded by a gene on chromosome 6p12-22 contains
CC RT extensive coiled-coil alpha-helical domains and a granin motif.";
CC RL J. Biol. Chem. 271:8328-8337(1996).
CC RN
CC RP SEQUENCE FROM N.A.
CC RA Seelig H.P.;
CC RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
CC RN
CC RP SEQUENCE OF 131-2230 FROM N.A.
CC RC TISSUE=Placenta;
```

RX MEDLINE=96125112; PubMed=8537393;
 RA Pritzler M.J., Lung C.-C., Hamel J.C., Griffith K.J., Chan E.K.L.;
 RT "Molecular characterization of Golgin-245, a novel Golgi complex
 RL protein containing a granin signature.";
 RN J. Biol. Chem. 270:31262-31268(1995).
 [4]
 RP SEQUENCE OF 524-672 FROM N.A.
 RC TISSUE=Gastric fundus;
 RA Balague C.;
 RL Thesis (1994), Instituto municipal de investigacion medica, Spain.
 CC -!- FUNCTION: May play a role in vesicular transport from the trans-
 CC Golgi.
 CC -!- SUBCELLULAR LOCATION: ASSOCIATED WITH THE CYTOPLASMIC FACE OF THE
 CC GOLGI MEMBRANE.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=4;
 CC Comment=Additional isoforms seem to exist;
 CC Name=1;
 CC IsoId=Q13439-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q13439-2; Sequence=VSP_004273;
 CC Name=3;
 CC IsoId=Q13439-3; Sequence=VSP_004274;
 CC Name=4;
 CC IsoId=Q13439-4; Sequence=VSP_004275;
 CC -!- DISEASE: ANTIGEN IN THE AUTOIMMUNE DISEASE SJOEGREN'S SYNDROME AND
 CC IN HEPATITIS B.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U41740; AAC50434.1; -;
 DR EMBL; X82834; CAA58041.1; -;
 DR EMBL; U31906; AAC51791.1; -;
 DR EMBL; X76942; CAA54261.1; -;
 DR Genbank; HGNC:4427; GOLGA4.
 DR MIM; 602509; -;
 DR GO; GO:0005802; C:Golgi trans-face; TAS.
 DR GO; GO:0016192; P:vesicle-mediated transport; TAS.
 DR InterPro; IPR000237; GRIP_domain.
 DR Pfam; PF01465; GRIP; 1.
 KW Golgi stack; Antigen; Coiled coil; Alternative splicing.
 FT DOMAIN 133 237
 FT COILED COIL (POTENTIAL).
 FT DOMAIN 276 1011
 FT COILED COIL (POTENTIAL).
 FT DOMAIN 1033 1214
 FT COILED COIL (POTENTIAL).
 FT DOMAIN 1259 2152
 FT TPYKGNLYHDTVSLFGFPTFEYLKVLPEY -> HLTKV
 FT VARSPPLIC 2154 2185
 FT AICTIRMGSHLENPLNSICKCPULSI (in isoform
 FT 2).
 FT /FTId=VSP_004272.
 FT Missing (in isoform 2).
 FT /FTId=VSP_004273.
 FT Missing (in isoform 3).
 FT /FTId=VSP_004274.
 FT /FTId=VSP_004275.
 FT FTSPRSGLF -> SWLRSSS (in isoform 4).
 FT /FTId=VSP_004275.
 FT R -> K (IN REF. 3).
 FT Y -> H (IN REF. 3).
 FT T -> A (IN REF. 3).
 FT K -> E (IN REF. 3).
 FT T -> A (IN REF. 3).
 FT K -> E (IN REF. 3).
 FT G30 630
 FT K -> N (IN REF. 3).
 FT CONFLICT 682 682
 FT 2330 AA; 261139 MW; 3BB733DB1EA86134 CRC64;
 SQ SEQUENCE
 Query Match 13.2%; Score 87.5; DB 1; Length 2230;
 Best Local Similarity 25.7%; Pred. No. 45;
 Matches 35; Conservative 29; Mismatches 45; Indels 27; Gaps 5;

Qy 5 IAPFVATLVAVTSGEDKKDYQNEFDPLMERIHEQIKKGELALFYLOQOINFEKPT 64
 Db 384 IQLRSRIKQMTTQGES-----LREQEKSERAAFBELEKALSTAK-T 426
 Qy 65 KEVKDKIVAEEMDTIAMIADGVGLDLMORKDLDFEQYNLEMLKSGD-----IL 116
 Db 427 EEARLKLAEMDEQIKTIKT-SEERISLQOELSRVKQEVVDVVMKKSSEBQIAKLOKLH 485
 Qy 117 ERDL-KKEEARVKKIE 131
 Db 486 EKELARKEQELTKLQ 501
 RESULT 7
 PIBF HUMAN
 ID PIBF HUMAN STANDARD; PRT; 757 AA.
 AC Q8XW3; O95664; Q96SF4;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Progesterone-induced blocking factor 1.
 GN PIBF1 OR PIBF.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CC NCBI_TaxID=9606;
 RN [1]
 RN SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RC TISSUE=Uterus;
 RX MEDLINE=21932550; PubMed=11935316;
 RA Rozenblum E., Vahteristo P., Sandberg T., Berghorsson J.T.,
 RA Syrjakoski K., Weaver D., Haraldsson K., Johannsdottir H.K.,
 RA Vermanen P., Ngam S., Golberger N., Robbins C., Pak E., Dutra A.,
 RA Gillander E., Stephan D.A., Bailey-Wilson J., Juo S.H., Kainu T.,
 RA Arason A., Barkardottir R.B., Nevanlinna H., Borg A.,
 RA Kallioniemi O.P.;
 RA "A genomic map of a 6-Mb region at 13q21-q22 implicated in cancer
 RT development: identification and characterization of candidate
 RT genes.";
 RL Hum. Genet. 110:111-121(2002).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Szekeres-Bartho J.;
 RN Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN SEQUENCE OF 1-611 FROM N.A.
 RA Blakey S.;
 RN Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP REVIEW.
 RX MEDLINE=21300944; PubMed=11407300;
 RA Szekeres-Bartho J., Barakonyi A., Par G., Polgar B., Palkovics T.,
 RA Szerecsenyi L.;
 RT "Progesterone as an immunomodulatory molecule.";
 RL Int. Immunopharmacol. 1:1037-1048(2001).
 CC -!- FUNCTION: Mediator of progesterone that by acting on the
 CC phospholipase A2 enzyme interferes with arachidonic acid
 CC metabolism, induces a Th2 biased immune response, and by
 CC controlling NK activity exerts an anti-abortion effect.
 CC -!- TISSUE SPECIFICITY: Expressed at highest levels in testis.
 CC Moderate expression is detected in spleen, thymus, prostate,
 CC ovary, small intestine, and colon.
 CC -!- INDUCTION: By progesterone.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```

CC -----
DR EMBL; AF330046; AAL37481.1;
DR EMBL; Y09631; CAA70844.1;
DR EMBL; AL391384; CAC39624.1;
DR MW; 607532;
FT CONFLICT 155 N -> S (IN REF. 2).
FT CONFLICT 333 R -> C (IN REF. 2).
FT CONFLICT 358 MYEKVAVSR -> I (IN REF. 3).
FT CONFLICT 597 DLEHRKQVLTQSPSLDR -> RSTGSKGPSNTAFTRSLTE
      (IN REF. 2)
FT CONFLICT 610 P -> Q (IN REF. 3).
FT CONFLICT 757 AA; 89773 MW; B468EB2A071ESB8D CRC64;
SQ SEQUENCE 757 AA; 89773 MW; B468EB2A071ESB8D CRC64;

Query Match 13.1%; Score 87; DB 1; Length 757;
Best Local Similarity 23.8%; Pred. No. 16;
Matches 29; Conservative 34; Mismatches 39; Indels 20; Gaps 5;

QY 12 LAVMTVSGEDKHHYQNFDFLLMERIHEQIKKKGELALFYLOEQNHFEKPTKEMDKI 71
Db 67 LSQKTMIDNLKVDYLTKE-ELEKLNDAHFQKQLLTLLRLDNQL-AFOQKDASKYQELM 124
QY 72 VAEQMTTIAMIDGVRGLDRMQRKDLDFQYNLEMLKSGDILE--RDLKKEARVKK 129
Db 125 KQEMETI-----LHQKQL---ETNLQLEKAGDVRRNLRDFELTEEOYIK 168
QY 130 IE 131
Db 169 LK 170

RESULT 8
RA50_AQUAE STANDARD; PRT; 978 AA.
AC O67124;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable DNA double-strand break repair rad50 ATPase.
GN RAD50 OR AQ 1006.
OS Aquifex aeolicus.
OC Bacteria; Aquificae.
OC Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus."
RL Nature 392:353-358(1998).
CC -!- FUNCTION: Involved in DNA double-strand break repair (DSBR). The
CC rad50/mre11 complex possesses single-strand endonuclease activity
CC and ATP-dependent double-strand-specific exonuclease activity.
CC Rad50 provides an ATP-dependent control of mre11 by unwinding
CC and/or positioning DNA ends into the mre11 active site (By
CC similarity).
CC -!- SUBUNIT: Forms a complex with mre11 (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000718; AAC07092.1;
DR F01; A70387; A70387.
DR HAMAP; MF_00449; -.
DR InterPro; IPR003439; ABC_transporter.

```

```

DR InterPro; IPR004592; SbcC.
DR InterPro; IPR002017; Spectrin.
DR Pfam; PF04430; DUF498; 1.
DR TIGRFAMs; TIGR00618; sbcc; 1.
KW DNA repair; Hydrolase; ATP-binding; Coiled coil; Complete proteome.
FT NP_BIND 32 39 ATP (BY SIMILARITY).
FT DOMAIN 160 826 COILED COIL (POTENTIAL).
SQ SEQUENCE 978 AA; 115897 MW; 9B0F2BF51ADD1151 CRC64;

Query Match 13.1%; Score 86.5; DB 1; Length 978;
Best Local Similarity 24.4%; Pred. No. 23;
Matches 30; Conservative 33; Mismatches 41; Indels 19; Gaps 5;

QY 12 LAVMTVSGEDKHHYQNFDFLLMERIHEQIKKKGELALFYLOEQNHFEKPTKEMDKI 121
Db 355 LSQSSLSLKEKREYEQAKQEF-----EDLSERVEKGGKLVAEETEEKIKELFSEEEY 409
QY 63 PTKEMDKIYVAEMDTTIAMIDGVRGLDRMQRKDLDFQYNLEMLKSGDILERDLK 121
Db 410 TSLMKERLLVELQKELKEGQLENLTQKYKKEKKVHKVNLKE-----LERELK 465
QY 122 KEE 124
Db 466 ERE 468

RESULT 9
PMW_DROME STANDARD; PRT; 254 AA.
AC Q9VTZ6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable phosphomannomutase (EC 5.4.2.8) (PMW).
GN CGI10688.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Ffannkoch C., Baldwin D.,
RA Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Bahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hestlin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler P., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

```



```

DR PIR; S73342; S73342.
DR InterPro; IPR002862; DUF16.
DR Pfam; PF01519; DUF16; 1.
DR KW Hypothetical protein; Complete proteome.
DR SQ SEQUENCE 166 AA; 19518 MW; BE44F5377B2FA709 CRC64;

Query Match 12.7%; Score 84; DB 1; Length 166;
Best Local Similarity 21.1%; Pred. No. 5.8;
Matches 28; Conservative 30; Mismatches 57; Indels 18; Gaps 3

QY 17 VSGEDKKHYQNEF-----DPLMERIHEQIKGELALFYIQEINHPFE 61
DB 31 INPENKVFESGFNFRQKLNKPIVYNTVEYTKFENYKDSNDQRLTKIENKVKDKLEV 90
QY 62 KPTK--EMDKIVAEIMDTIIAMIDGVGVGLRLMQKRD-LDIFQYNLEMLKKSGLIER 118
DB 91 KVKLEKVDKLEAKVDKLEEKVDKLEAKVDKLEAKVDKLEAKVDKLEAKVDKLEAKVDK 150
QY 119 DLKKEEARVKIE 131
DB 151 RLDSIEGLDKIE 163

RESULT 13
UVR_C BORBU STANDARD; PRT; 603 AA.
AC OS1413;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE UVRABC system protein C (UvrC protein) (Excinuclease ABC subunit C).
DE UVR_C OR BB0457.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Karlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Uterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.";
RL Nature 390:580-586(1997).
CC -!- FUNCTION: The UvrABC repair system catalyzes the recognition and
CC processing of DNA lesions. UvrC both incises the 5' and 3' sides
CC of the lesion. The N-terminal half is responsible for the 3'
CC incision and the C-terminal half is responsible for the 5'
CC incision (By similarity).
CC -!- SUBUNIT: Interacts with uvrB in an incision complex (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the uvrC family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement
or send an email to license@isb-sib.ch).

EMBL; AE001150; AAC66813.1; -.
PIR; H70156; H70156.
TIGR; BB0457; -.
DR HAMAP; MF 00203; -. 1.
DR InterPro; IPR004791; UvrC.

```

DR InterPro; IPR001162; UvrC C.
DR InterPro; IPR000305; UvrC N.
DR Pfam; PF01541; Exci_endo_N.
DR ProDom; PD003870; UvrC_C; 1.
DR SMART; SM00465; GIYC; 1.
DR TIGRFAMs; TIGR00194; UvrC; 1.
DR PROSITE; PS50164; UVR_C; 1.
DR PROSITE; PS50165; UVR_C2; 1.
DR SOS response; Excision nuclease; DNA repair; DNA recombination;
KW DNA excision; Complete procenome.
SQ SEQUENCE 603 AA; 69832 MW; B3EDDB6B4A63484 CRC64;

Query Match 12.6%; Score 83.5; DB 1; Length 603;
Best Local Similarity 21.1%; Pred. No. 24;
Matches 30; Conservative 28; Mismatches 53; Indels 31; Gaps 5;

Qy 1 MKFTIAFFVATLAVMTVSGEDKHDYQNEPFLMERIHBOIKKGELALFYLBQ----- 55
Db 291 LQFLIQVYTSINMIV---PDKIHFLKIDTKNVEKLINEIKNTKTEIYKETEELIKI 346

Qy 56 -----INHFEPKPTKEMKD-KIVAEMDII-----IAMDVGVRGVLRLMQR 95
Db 347 MEMAISNAELSLRYENKSTKALESLKIVLEMDKLPKTIIEGFDIAHLKGQETVASMVTFK 406

Qy 96 KDLDFEFOYNLEMLKK--SGDI 115
Db 407 MGMPFKNRYRLKLSLLKGEI 428

RESULT 14
MYSJ DICDI STANDARD; PRT; 2245 AA.
AC P54697;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Myosin IJ heavy chain.
GN MYOJ.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
ON NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX3;
RX MEDLINE=96215148; PubMed=8636147;
RA Hammer J.A. III, Jung G.;
RT "The sequence of the dictyostelium myo J heavy chain gene predicts a
RT novel, dimeric, unconventional myosin with a heavy chain molecular
RT mass of 258 kDa."
RL J. Biol. Chem. 271:7120-7127(1996).
RN [2]
RP SEQUENCE OF 1-1021 FROM N.A.
RX MEDLINE=97039016; PubMed=8884597;
RA Peterson M.D., Urioste A.S., Titus M.A.;
RT "Dictyostelium discoideum myoJ: a member of a broadly defined myosin
RT V class or a class XI unconventional myosin?";
RL J. Muscle Res. Cell Motil. 17:411-424(1996).
RN [3]
RP SEQUENCE OF 182-298 FROM N.A.
RX MEDLINE=95023928; PubMed=7937787;
RA Titus M.A., Kuspa A., Loomis W.F.;
RT "Discovery of myosin genes by physical mapping in Dictyostelium."
RL Proc. Natl. Acad. Sci. U.S.A. 91:9446-9450(1994).
CC -!- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN AND HAS ATPASE
CC ACTIVITY THAT IS ACTIVATED BY ACTIN.
CC -!- SUBUNIT: Homodimer.
CC -!- SIMILARITY: Contains 1 myosin-like globular head domain.
CC -!- SIMILARITY: Contains 1 dilute domain.
CC -!- SIMILARITY: Contains 3 IQ domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U42409; AAA85186.1; -;
DR EMBL; L35322; AAA79858.1; -;
DR PIR; T18278; T18278.
DR HSP; P08799; INWD.
DR DictyDb; DD01095; myoJ.
DR InterPro; IPR002710; DIL.
DR InterPro; IPR000048; IQ region.
DR InterPro; IPR001609; myosin head.
DR InterPro; IPR004009; Myosin_N.
DR Pfam; PF01843; DIL; 1.
DR Pfam; PF00612; IQ; 6.
DR Pfam; PF00063; myosin head; 2.
DR Pfam; PF02736; Myosin_N; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD003376; DIL; 1.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 3.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 3.
KW Myosin; ATP-binding; Actin-binding; Calmodulin-binding; Repeat;
KW Coiled coil.
FT DOMAIN 1 809 MYOSIN HEAD-LIKE.
FT DOMAIN 824 851 IQ 1.
FT DOMAIN 872 901 IQ 2.
FT DOMAIN 943 972 IQ 3.
FT DOMAIN 973 1812 COILED COIL (POTENTIAL).
FT DOMAIN 2060 2220 DILUTE.
FT NP_BIND 174 181 ATP (POTENTIAL).
FT DOMAIN 669 749 ACTIN-BINDING.
FT CONFLICT 191 191 L -> F (IN REF. 2).
FT CONFLICT 284 284 A -> T (IN REF. 2).
FT CONFLICT 291 291 G -> R (IN REF. 2).
FT CONFLICT 332 347 NKSGCPRIEIGVSDEEH -> IEMFELKLVYRKMS (IN
FT CONFLICT 550 550 N -> K (IN REF. 2).
FT CONFLICT 865 866 HH -> QQ (IN REF. 2).
SQ SEQUENCE 2245 AA; 258478 MW; 615E5EF1D1AB45BE CRC64;

Query Match 12.5%; Score 83; DB 1; Length 2245;
Best Local Similarity 27.4%; Pred. No. 1e+02;
Matches 23; Conservative 17; Mismatches 34; Indels 10; Gaps 3;

Qy 39 HEQIKKGELALFYLBQEQINHFEE--KPTKEMKQKIVAEMDTIAMDVGVRGVLRLMQRK 96
Db 1082 HDSIEK-----LQSQFNETEQLOQFKQSGSELSKLSKTTTQQLDFNKQDFRLSQER 1134

Qy 97 DLD-IFEQYNLEMLKSGDILERD 119
Db 1135 DTDNTNNQLEIQQLKKANSTLEED 1158

RESULT 15
SYA THEIN STANDARD; PRT; 878 AA.
AC Q8RAH4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alanyl-tRNA synthetase (EC 6.1.1.7) (Alanine--tRNA ligase) (AlaRS).
GN ALAS OR TTE1248.
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacterales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336;

```

RA  Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA  Chen X., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA  Tan H., Chen R., Wang J., Yu J., Yang H.;
RT  "A complete sequence of T. tengcongensis genome.";
RL  Genome Res. 12:689-700(2002).
CC  -!- CATALYTIC ACTIVITY: ATP + L-alanine + tRNA(Ala) = AMP +
CC  diphosphate + L-alanyl-tRNA(Ala).
CC  -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC  -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AE013087; AAM24472.1; -.
DR  HAMAP; MF 00036; -; 1.
DR  InterPro; IPR003156; DHHA1.
DR  InterPro; IPR002318; tRNA-synt_2c.
DR  InterPro; IPR006193; tRNA_synt_Ala.
DR  Pfam; PF02272; DHHA1; 1.
DR  Pfam; PF01411; tRNA-synt 2c; 1.
DR  TIGRFAMs; TIGR00344; alaS; 1.
DR  PROSITE; PS50860; AA tRNA LIGASE II ALA; 1.
KW  Aminoacyl-tRNA synthetase; Protein Biosynthesis; Ligase; ATP-binding;
KW  Complete proteome.
SQ  SEQUENCE 878 AA; 98363 MW; F2CE98288DC2A244 CRC64;
Query Match
Best Local Similarity 29.8%; Pred No.42;
Matches 31; Conservative 22; Mismatches 32; Indels 19; Gaps 6;
QY 26 YQNEFDLLMERIHQIKKGELALF-----YLQEQINHFEEKPTKEMKDKIVAEM-DTII- 79
Db 706 YLNEQKEIL-DRVSETLKAQKEVSVKTIENLQSLKD-KERIEGLKTKIASILAETLID 763
QY 80 --AMIDGVRGVLDRMQRKDLDFEQYNLEMLKKSGLIERDLK 121
Db 764 SAISVDGVKVIASRV-----EDYDMEALKTLGLDKRLK 798

```

Search completed: November 17, 2003, 08:23:39
 Job time : 18 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2003, 08:21:04 ; Search time 34 Seconds
(without alignments)
1001.851 Million cell updates/sec

Title: US-09-877-160A-1

Perfect score: 662

Sequence: 1 MKFIIAFVATLAVMTSGE.....GDILERDLKKEARVKIEV 132

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SEPREMBL_23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mmc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	103	15.6	2138	5 Q9XZE3	Q9xe3 amoeba prot
2	94.5	14.3	506	16 Q8E613	Q8e613 streptococc
3	94.5	14.3	506	16 Q8E128	Q8e128 streptococc
4	94	14.2	860	13 Q42351	Q42351 gallus gall
5	92	13.9	776	16 Q8XK4	Q8xk4 clostridium
6	90.5	13.7	920	10 Q8S4W6	Q8s4w6 chlamydomon
7	89	13.4	249	16 Q9ZK05	Q9zk05 helicobacte
8	88.5	13.4	558	16 Q8RGW3	Q8rgw3 fusobacteri
9	88.5	13.4	846	4 O75130	O75130 homo sapien
10	88.5	13.4	939	16 Q98Q18	Q98q18 mycoplasma
11	87.5	13.2	459	5 Q04011	Q04011 wuchereria
12	87.5	13.2	623	4 Q9H2G3	Q9h2g3 homo sapien
13	87.5	13.2	2228	5 Q9I1W5	Q9i1w5 plasmodium
14	87	13.1	311	17 Q971V5	Q971v5 sulfolobus
15	86	13.0	249	16 Q25166	Q25166 helicobacte
16	86	13.0	301	16 Q8XL82	Q8xl82 clostridium

RESULT 1

Q9XZE3 PRELIMINARY; PRT; 2138 AA.
 ID Q9XZE3
 AC Q9XZE3;
 DT 01-NOV-1999 (TRENBLrel. 12, Created)
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Myosin heavy chain.
 OS Amoeba proteus (Amoeba).
 OC Eukaryota; Lobosea; Euamoebida; Amoebidae; Amoeba.
 OX NCBI_TaxID=5775;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=D;
 RX MEDLINE=99082369; PubMed=9864850;
 RA Oh S.W., Jeon K.W.;
 RT "Characterization of myosin heavy chain and its gene in Amoeba proteus.";
 RL J. Eukaryot. Microbiol. 45:600-605(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=D;
 RA Oh S.W., Jeon K.W.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 EMBL; AF136711; AAD33718.1; -.
 DR HSSP; P08799; 1MND.
 DR InterPro; IPR000048; IQ_region.
 DR InterPro; IPR001609; myosin_head.
 DR InterPro; IPR004009; Myosin_N.
 DR Pfam; PF00612; IQ; 2.
 DR Pfam; PF02736; myosin_head; 1.
 DR Pfam; PF00063; myosin_N; 1.
 DR PRINTS; PR00193; MYOSINHEAVY.
 DR ProDom; PD000355; myosin_head; 1.
 DR SMART; SM00015; IQ; 1.
 DR SMART; SM00242; MYSC; 1.
 DR PROSITE; PS00096; IQ; 1.
 SQ SEQUENCE 2138 AA; 244427 MW; D58B8308639FA71 CRC64;

ALIGNMENTS

17 86 13.0 1158 16 Q8XPE3
 18 86 13.0 3187 11 Q83714
 19 85 12.8 230 2 Q8GL35
 20 85 12.8 784 16 Q8ZRB9
 21 85 12.8 784 16 Q8ZBVO
 22 85 12.8 784 16 Q8XE60
 23 85 12.8 2413 5 Q8ICH7
 24 84.5 12.8 536 12 Q8JUS1
 25 84 12.7 185 16 Q51736
 26 84 12.7 1650 5 Q812T7
 27 83.5 12.6 215 16 Q8EW76
 28 83.5 12.6 311 6 Q95K18
 29 83 12.5 155 16 Q8EMU6
 30 83 12.5 559 5 Q02263
 31 83 12.5 951 5 Q96229
 32 83 12.5 1310 10 Q949K0
 33 82.5 12.5 233 2 Q44739
 34 82.5 12.5 445 16 Q9PRJ8
 35 82.5 12.5 476 6 Q8WP20
 36 82.5 12.5 1201 16 Q8DAL0
 37 82 12.4 597 5 Q9NCN3
 38 82 12.4 599 5 Q9NCM5
 39 82 12.4 599 5 Q9NCM7
 40 82 12.4 690 5 Q8SSE8
 41 82 12.4 1558 5 Q96275
 42 82 12.4 1716 5 Q8IIM2
 43 82 12.4 1751 5 Q26194
 44 82 12.4 1756 5 Q8IIM1
 45 82 12.4 1756 5 Q8I058

Q8xpe3 clostridium
 Q63714 rattus norv
 Q8gl35 borrelia bu
 Q8zrb9 salmonella
 Q8z8v0 salmonella
 Q8xe60 escherichia
 Q8ich7 plasmodium
 Q8js31 phthorimaea
 Q51736 borrelia bu
 Q812T7 plasmodium
 Q8ew76 mycoplasma
 Q95k18 macaca fasc
 Q8emu6 oceanobacil
 Q02263 caenorhabdi
 Q96229 plasmodium
 Q949k0 lycopersico
 Q44739 borrelia bu
 Q9ppj8 campylobact
 Q8wp20 macaca fasc
 Q8dal0 vibrio vuln
 Q9ncn3 plasmodium
 Q9ncm5 plasmodium
 Q9ncm7 plasmodium
 Q8sse8 encephalico
 Q96275 plasmodium
 Q8iim2 plasmodium
 Q26194 plasmodium
 Q8iim1 plasmodium
 Q8i058 plasmodium

```
Query Match      15.6%; Score 103; DB 5; Length 2138;
Best Local Similarity 29.6%; Pred. No. 7.1;
Matches 29; Conservative 24; Mismatches 35; Indels 10; Gaps 3;

QY 35 MERIHQIKKGELALFYLOQINHFEEKPTKEMKDKIVAEMDTIAMDGVGVLDRLMQ 94
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1730 LBLEEQLEAEAEALNYL-BEIKHKDLLEINELRKQLDAESE-----ARDKFEQLKN 1780

QY 95 RKDLDFE-CYNLEMLKSGDILERDLKKEARVKIE 131
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1781 ELERDVADAHNLEAEKSKTDAERAKKAQYDELK 1818

RESULT 2
Q8B613 PRELIMINARY; PRT; 506 AA.
AC Q8B613;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DN Hypothetical protein.
GN GBS0580.
OS Streptococcus agalactiae (serotype III).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216495;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEM316 / Serotype III;
RX MEDLINE=22242508; PubMed=12354221;
RA Glaser P., Rusniok C., Buchrieser C., Chevalier F., Frangeul L.,
RA Msadek T., Zouine M., Couve E., Lalioui L., Poyart C., Trieu-Cuot P.,
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
RT invasive neonatal disease.";
RL Mol. Microbiol. 45:1499-1513 (2002).
DR EMBL; AL766846; CAD46224.1; -.
DR Sgallist; gbs0580; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 506 AA; 57621 MW; A617EAC632CF8CDB CRC64;

Query Match      14.3%; Score 94.5; DB 16; Length 506;
Best Local Similarity 26.3%; Pred. No. 7.7;
Matches 25; Conservative 24; Mismatches 33; Indels 13; Gaps 3;

QY 45 GELALFYLOQIN--HFEEKPTKEMKDKIVAEMDTIAMDGVGVLDRLMQKDLDFE 102
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 240 GELAKYIKKYNINYYFEENASKNVAKTLADEVGVKTAVLSPLEG-----LSKEMAAGE 294

QY 103 QY-----NLEMLKSGDILERDLKKEARVKIE 131
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 295 DYFSVMERNLKVLTDTDVAGKEVAPEEDKTKIVE 329

RESULT 3
Q8E128 PRELIMINARY; PRT; 506 AA.
AC Q8E128;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Zinc ABC transporter, zinc-binding adhesion liprotein.
GN SAG0535.
OS Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2603 V/R / Serotype V;
RX MEDLINE=2222988; PubMed=12200547;
RA Tettelin H., Masiagnani V., Cieslewicz M.J., Eisen J.A., Peterson S.,
RA Wessels-W.R., Paulsen I.T., Nelson K.E., Margalit I., Read T.D.,
```

```
RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA Desoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,
RA Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Moza M.,
RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni P., Maione D.,
RA Rinaldo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an
RT emerging human pathogen, serotype V Streptococcus agalactiae";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396 (2002).
DR EMBL; AB014216; AAM99436.1; -.
DR TIGR; SAG0535; -.
KW Complete proteome.
SQ SEQUENCE 506 AA; 57621 MW; A617EAC632CF8CDB CRC64;

Query Match      14.3%; Score 94.5; DB 16; Length 506;
Best Local Similarity 26.3%; Pred. No. 7.7;
Matches 25; Conservative 24; Mismatches 33; Indels 13; Gaps 3;

QY 45 GELALFYLOQIN--HFEEKPTKEMKDKIVAEMDTIAMDGVGVLDRLMQKDLDFE 102
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 240 GELAKYIKKYNINYYFEENASKNVAKTLADEVGVKTAVLSPLEG-----LSKEMAAGE 294

QY 103 QY-----NLEMLKSGDILERDLKKEARVKIE 131
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 295 DYFSVMERNLKVLTDTDVAGKEVAPEEDKTKIVE 329

RESULT 4
O42351 PRELIMINARY; PRT; 860 AA.
AC O42351;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE Chicken RABAPTIN-5 (NEUROCRESCIN).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RC TISSUE=Brain;
RA Nogawa M., Uyeda A., Nishimune H., Taguchi T.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DDJ databases.
RN [2]
RP SEQUENCE OF 79-860 FROM N.A.
RC TISSUE=Muscle, and Telencephalon;
RX MEDLINE=98087215; PubMed=9427343;
RA Nishimune H., Uyeda A., Nogawa M., Fujimori K., Taguchi T.;
RT "Neurorescin: a novel neurite-outgrowth factor secreted by muscle
RT after denervation.";
RL NeuroReport 8:3649-3654 (1997).
DR EMBL; D88828; BAA21785.1; -.
DR InterPro; IPR003914; Rabaptin.
DR Pfam; PF03528; Rabaptin; 2.
DR PRINTS; PR01432; RABAPTIN.
SQ SEQUENCE 860 AA; 99414 MW; 202FCF2CA845B2BE CRC64;

Query Match      14.2%; Score 94; DB 13; Length 860;
Best Local Similarity 27.9%; Pred. No. 15;
Matches 41; Conservative 26; Mismatches 46; Indels 34; Gaps 8;

QY 16 TVSGEDKKH-DYQNEFDLLMERIHEQIKKGELALFYLOEQI-----NHFEKPTKE-- 66
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 649 SLQGHSLVSLQQAEDFILPEAE-----LRELILKYREDIISVTAADHLEKLAEL 705

QY 67 MKDKIVAE-----DTI-----IAMDGVGVLDRL--MQKDLDFE---Y 104
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 706 FLKEQIQAEQVLYQNIETLQLEIENCKEEIASISLAKAELEIKVEKEQLESSQEVILQ 765
```

```

QY 105 NLEMLKSGDILRLDKKEARVKIE 131
  |||:::|||||:::
Db 766 QLESQETKNTLEBQLKKTAAKANLE 792

RESULT 5
Q8XKK4 PRELIMINARY; PRT; 776 AA.
AC Q8XKK4;
DT 01-WAR-2002 (TREMELrel. 20, Created)
DT 01-WAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-WAR-2003 (TREMELrel. 23, Last annotation update)
DE ATP-dependent protease La.
GN LONA OR CPE1390.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / Type A;
RX PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR EMBL; AP003190; BAB91096.1; -.
DR MEROPS; S16.001; -.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR003959; AAA ATPase_cent.
DR InterPro; IPR003111; LON.
DR InterPro; IPR001984; Lon_endopep.
DR InterPro; IPR004815; Lon_fam.
DR Pfam; PF00004; AAA; 1.
DR Pfam; PF02190; LON; 1.
DR PRINTS; PR00830; ENDOLAPTASE.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00464; LON; 1.
DR TIGRFAMs; TIGR00763; lon; 1.
DR PROSITE; PS01046; LON_SER; 1.
KW Protease; Complete proteome.
SQ SEQUENCE 776 AA; 87428 MW; F818519BF7174D19 CRC64;

Query Match 13.9%; Score 92; DB 16; Length 776;
Best Local Similarity 27.7%; Pred. No. 19;
Matches 41; Conservative 21; Mismatches 58; Indels 28; Gaps 7;

QY 1 MKFIIAFVATLAVMTVSGEDKKHDYQNEFDPLMERIHEQIKKGELALFYIQEINHFE 60
  |||:::|||||:::
Db 74 IKQIVKMPKNTIRVLVEGIERAKMD--BFFD-----KEELLEASIEKIDIDNEIDHEL 124
  |||:::|||||:::

QY 61 EKPTEKMDKIVAEMDTIAMDVGVRV--LDRIMQRDL-----DIFQYNLEMLKKSQD 114
  |||:::|||||:::
Db 125 BALSRKLKDDPFELD--ITASSGNGVDFLNLESEKDLNKTVDLISYALIKQEDKQD 182
  |||:::|||||:::

QY 115 ILER-DLK-----KEARVKIE 131
  |||:::|||||:::
Db 183 ILQTLDLKRIEKLIIFYVKEIEVAKIE 210
  |||:::|||||:::

RESULT 6
Q8S4W6 PRELIMINARY; PRT; 920 AA.
AC Q8S4W6;
DT 01-JUN-2002 (TREMELrel. 21, Created)
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)
DE Coiled-coil flagellar protein.
GN MBO2.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.

```

```

OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11977094;
RA Tam L.W., Lefebvre P.A.;
RT "The Chlamydomonas MBO2 locus encodes a conserved coiled-coil protein
RT important for flagellar waveform conversion.";
RL Cell Motil. Cytoskeleton 51:197-212(2002).
DR EMBL; AF394181; AAM15771.1; -.
KW Flagella.
SQ SEQUENCE 920 AA; 104210 MW; 9B24A671A6408830 CRC64;

Query Match 13.7%; Score 90.5; DB 10; Length 920;
Best Local Similarity 23.9%; Pred. No. 31;
Matches 38; Conservative 30; Mismatches 38; Indels 53; Gaps 6;

QY 16 TVSGEDKKHDYQNEFDPL-----LMERIHEQ-----IKKGELALFYIQE 54
  |||:::|||||:::
Db 509 TTEMKDKLVQLQNELDILQNEVGIKDKLLQOCHTOHAANIAERDQLRVELGELGNVFRDK 568
  |||:::|||||:::

QY 55 QINHEPEKPTKEMDKIVAEMDTIAMDVGVRVLDRLMQRKOLDI----- 100
  |||:::|||||:::
Db 569 Q-----AVVDEQIARVDKLNAINGCEKEMLRKQVELVEARNVTGIMLDNRN 618
  |||:::|||||:::

QY 101 -----PROYN-LFMLKSGDILRLDKKEARVKIEV 132
  |||:::|||||:::
Db 619 DELCVLYEKANILDEVKSGQ-LELMRREDEARLRLLEV 656
  |||:::|||||:::

RESULT 7
Q9ZKG5 PRELIMINARY; PRT; 249 AA.
AC Q9ZKG5;
DT 01-MAY-1999 (TREMELrel. 10, Created)
DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)
DT 01-MAY-2000 (TREMELrel. 13, Last annotation update)
DE Putative PARALOG of HPAa.
GN JHP0971.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=99236862;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
DR EMBL; AE001526; AAO06549.1; -.
KW Complete proteome.
SQ SEQUENCE 249 AA; 28513 MW; EC8D9AFDC99548AA CRC64;

Query Match 13.4%; Score 89; DB 16; Length 249;
Best Local Similarity 28.8%; Pred. No. 10;
Matches 34; Conservative 27; Mismatches 39; Indels 18; Gaps 8;

QY 25 DYQNEFD---FLMERIHEQIKKGELALFYIQEINHEPEKTKEMDKK--IIVEMDTII 79
  |||:::|||||:::
Db 76 NYQKEFKSLFLQLSNFLE--RKG-----YSVSQFKDVSEIP-QDIKEKALLVLRMDGNV 127
  |||:::|||||:::

QY 80 AMIDGVRGVLDRLMQRKOLDIFEQY-NLEMLK-KSGDILER---DLKKEARVKIEV 132
  |||:::|||||:::
Db 128 ALLEDIVESDALSEEKVIDMSSGYNLNFVEPKSEDIHISFGIDVSKIAVERVEL 185
  |||:::|||||:::

RESULT 8
Q8RGM3 PRELIMINARY; PRT; 558 AA.
ID Q8RGM3

```


01-DEC-2001 (TREMELrel. 19, Last annotation update)
 Myosin-like protein (Fragment).
 GN MYOSIN-LIKE PROTEIN.
 OS Wuchereria bancrofti.
 OC Eukaryota; Metazoa; Nematozoa; Chromadorea; Spirurida; Filarioidea;
 ON Onchocercidae; Wuchereria.
 OX NCBI_TaxID=6293;
 RN [1]
 RQ SEQUENCE FROM N.A.
 RP MEDLINE=91312398; PubMed=1857386;
 RX Raghavan N., McReynolds L.A., Maina C.V., Feinstein S.M.,
 RA Jayaraman K., Ottosen E.A., Numan T.B.;
 RT "A recombinant clone of Wuchereria bancrofti with DNA specificity for
 RL human lymphatic filarial parasites."
 RL Mol. Biochem. Parasitol. 47:63-71(1991).
 DR EMBL; M38213; AAA63560.1; -;
 FT NON TER 1
 SQ SEQUENCE 459 AA; 53314 MW; 5E4683958E9A9272B CRC64;
 Query Match 13.2%; Score 87.5; DB 5; Length 459;
 Best Local Similarity 25.4%; Pred. No. 26;
 Matches 32; Conservative 30; Mismatches 43; Indels 21; Gaps 6;
 QY 21 DKKHQYQNEFDLMLERIEHQI---KKGELALFYLOQINHEEPKPKEMKDKIVAEMDT 77
 DB 41 DKLTDRYND---LSTRVNDKLNHAKOEQNVFDDIQQLDDMEQK-VDDFLNKYITSQDL 95
 QY 78 LIAM--IDGVRGVLDRLMQRKDLDFE-QYNLEMLKKGDI-----LERDLKKEE 124
 DB 96 LIAMEDVDQHSLLQDIPSAIENTECDLKLNKKADNKNKINLLVPLEKDTKQEQ 155
 QY 125 ARVKKI 130
 DB 156 KLMQNV 161
 RESULT 12
 QY9H2G3 PRELIMINARY; PRT; 623 AA.
 ID Q9H2G3
 AC Q9H2G3;
 DT 01-MAR-2001 (TREMELrel. 16, Created)
 DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
 DE CTCL tumor antigen se20-7 (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RQ SEQUENCE FROM N.A.
 RP TISSUE=Testis;
 RX MEDLINE=21143360; PubMed=11149944;
 RA Eichmuller S., Usener D., Dummer R., Stein A., Thiel D.,
 RA Schandendorf D.;
 RT "Serological detection of cutaneous T-cell lymphoma-associated
 RT antigen5".
 RL Proc. Natl. Acad. Sci. U.S.A. 98:629-634 (2001).
 DR EMBL; AF273047; AAG34907.1; -;
 DR InterPro; IPR001990; Granin.
 DR PROSITE; PS00422; GRANINS_1; 1.
 FT NON TER 623
 SQ SEQUENCE 623 AA; 73815 MW; BB2A5AF664B791ED CRC64;
 Query Match 13.2%; Score 87.5; DB 4; Length 623;
 Best Local Similarity 25.7%; Pred. No. 36;
 Matches 35; Conservative 29; Mismatches 45; Indels 27; Gaps 5;
 QY 5 IAFVATVAMTVSGEDKKHQQNEFDLMLERIEHQIKKGELALFYLOQINHEEPT 64
 DB 47 IAQLSRKQMTTQEE-----LREQEKSERAAFEELEKALSTAQK-T 89
 QY 65 KEMKDKIVAEMDTIAMIADGVGVLDRLMQRKDLDFEIQYNLEMLKSGD-----IL 116
 DB 116
 QY 90 EEARKRLKAEMDEQIKTEKT-SSEERISLQOELSRVQEVVDVNMKSSBEQIAKLQKLH 148
 QY 117 EEDL-KKEEARVKIE 131
 DB 149 EKELARKOEELTKLQ 164
 RESULT 13
 QY11W5 PRELIMINARY; PRT; 2228 AA.
 ID QY11W5
 AC QY11W5;
 DT 01-MAR-2003 (TREMELrel. 23, Created)
 DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
 DE Hypothetical protein (Fragment).
 GN PF030320C.
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=36329;
 RN [1]
 RQ SEQUENCE FROM N.A.
 RP MEDLINE=22255708; PubMed=12368867;
 RX Hall N., Pain A., Barriman M., Churcher C., Harris B., Harris D.,
 RA Mungall C., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
 RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
 RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
 RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
 RA Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
 RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
 RA Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,
 RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
 RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
 RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitz B.,
 RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
 RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
 RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
 RA Sulston J.E., Craig A., Newbold C., Barrell B.G;
 RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13."
 RL Nature 419:527-531(2002).
 DR EMBL; AL034557; CAD49158.1; -;
 KW Hypothetical protein.
 FT NON TER 1
 SQ SEQUENCE 2228 AA; 265493 MW; 83D9341248D3B0B0 CRC64;
 Query Match 13.2%; Score 87.5; DB 5; Length 2228;
 Best Local Similarity 21.2%; Pred. No. 1.4e+02;
 Matches 24; Conservative 38; Mismatches 44; Indels 7; Gaps 2;
 QY 20 EDKKHQYQNEFDLMLERIEHQIKKGELALFYLOQINHEEPTKEMKDKIVAEMDTII 79
 DB 605 ENEKNLNKQKQICMDQAKLEIQKKLSL-----EKSNFQMKNKIIEKLTNKEKIK 659
 QY 80 AMIDGVRGVLDRLMQRKDLDFEIQYNLEMLKSGDILERDLKKEEARVKIEV 132
 DB 660 LSLR--IGKLENRINKTAHDKFQSFNYVEKSDMKREKIRKDKIMEKIKM 710
 RESULT 14
 QY11V5 PRELIMINARY; PRT; 311 AA.
 ID QY11V5
 AC QY11V5;
 DT 01-DEC-2001 (TREMELrel. 19, Created)
 DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
 DE Hypothetical protein ST1273.
 GN ST1273.
 OS Sulfolobus tokodaii.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=111955;
 RN [1]
 RQ SEQUENCE FROM N.A.
 RP STRAIN=JCM 10545 / 7;
 RC

01-DEC-2001 (TREMELrel. 19, Last annotation update)
 Myosin-like protein (Fragment).
 GN MYOSIN-LIKE PROTEIN.
 OS Wuchereria bancrofti.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
 ON Onchocercidae; Wuchereria.
 OX NCBI_TaxID=6293;
 RN [1]
 RQ SEQUENCE FROM N.A.
 RP MEDLINE=91312398; PubMed=1857386;
 RX Raghavan N., McReynolds L.A., Maina C.V., Feinstein S.M.,
 RA Jayaraman K., Ottosen E.A., Numan T.B.;
 RT "A recombinant clone of Wuchereria bancrofti with DNA specificity for
 RT human lymphatic filarial parasites."
 RL Mol. Biochem. Parasitol. 47:63-71(1991).
 DR EMBL; M38213; AAA63560.1; -;
 FT NON TER 1
 SQ SEQUENCE 459 AA; 53314 MW; 5E468395E8A9272B CRC64;
 Query Match 13.2%; Score 87.5; DB 5; Length 459;
 Best Local Similarity 25.4%; Pred. No. 26;
 Matches 32; Conservative 30; Mismatches 43; Indels 21; Gaps 6;
 QY 21 DKKHQYQNEFDLMLERIEHQI---KKGELALFYLOEQINHEEPKPKEMKDKIVAEMDT 77
 DB 41 DKLTDRYND---LSTRVNDKLNHAKOEQNVFDDIQQLDDMEQK-VDDFLNKYITSQDL 95
 QY 78 LIAM--IDGVRGVLDRLMQRKDLDFE-QYNLEMLKKGDI-----LERDLKKEE 124
 DB 96 LIAMEDVDQHSLLQDIPSAIENTECDKLNLLKKADNKNKINLLVPLEKDTKQEQ 155
 QY 125 ARVKKI 130
 DB 156 KLMQNV 161
 RESULT 12
 QY9H2G3 PRELIMINARY; PRT; 623 AA.
 ID Q9H2G3
 AC Q9H2G3;
 DT 01-MAR-2001 (TREMELrel. 16, Created)
 DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
 DE CTCL tumor antigen se20-7 (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RQ SEQUENCE FROM N.A.
 RP TISSUE=Testis;
 RX MEDLINE=21143360; PubMed=11149944;
 RA Eichmuller S., Usener D., Dummer R., Stein A., Thiel D.,
 RA Schandendorf D.;
 RT "Serological detection of cutaneous T-cell lymphoma-associated
 RT antigen5".
 RL Proc. Natl. Acad. Sci. U.S.A. 98:629-634 (2001).
 DR EMBL; AF273047; AAG34907.1; -;
 DR InterPro; IPR001990; Granin.
 DR PROSITE; PS00422; GRANINS_1; 1.
 FT NON TER 623
 SQ SEQUENCE 623 AA; 73815 MW; BB2A5AF664B791ED CRC64;
 Query Match 13.2%; Score 87.5; DB 4; Length 623;
 Best Local Similarity 25.7%; Pred. No. 36;
 Matches 35; Conservative 29; Mismatches 45; Indels 27; Gaps 5;
 QY 5 IAFVATVAMTVSGEDKKHQQNEFDLMLERIEHQIKKGELALFYLOEQINHEEPT 64
 DB 47 IAQLSRKQMTTQEE-----LREQEKSERAAFEELEKALSTAQK-T 89
 QY 65 KEMKDKIVAEMDTIAMIADGVGVLDRLMQRKDLDFEYQINLEMLKSGD-----IL 116
 DB 116
 QY 90 EEARKKLAEMDEQIKTEKT-SSEERISLQOELSRVQEVVDVNMKSSBEQIAKLQKLH 148
 QY 117 EEDL-KKEEARVKKIE 131
 DB 149 EKELARKOEELTKLQ 164
 RESULT 13
 QY11W5 PRELIMINARY; PRT; 2228 AA.
 ID QY11W5
 AC QY11W5;
 DT 01-MAR-2003 (TREMELrel. 23, Created)
 DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
 DE Hypothetical protein (Fragment).
 GN PF030320C.
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=36329;
 RN [1]
 RQ SEQUENCE FROM N.A.
 RP MEDLINE=22255708; PubMed=12368867;
 RX Hall N., Pain A., Barriman M., Churcher C., Harris B., Harris D.,
 RA Mungall C., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
 RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
 RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
 RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
 RA Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
 RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
 RA Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,
 RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
 RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
 RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitz B.,
 RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
 RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
 RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
 RA Sulston J.E., Craig A., Newbold C., Barrell B.G;
 RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13."
 RL Nature 419:527-531(2002).
 DR EMBL; AL034557; CAD49158.1; -;
 KW Hypothetical protein.
 FT NON TER 1
 SQ SEQUENCE 2228 AA; 265493 MW; 83D9341248D3B0B0 CRC64;
 Query Match 13.2%; Score 87.5; DB 5; Length 2228;
 Best Local Similarity 21.2%; Pred. No. 1.4e+02;
 Matches 24; Conservative 38; Mismatches 44; Indels 7; Gaps 2;
 QY 20 EDKKHQYQNEFDLMLERIEHQIKKGELALFYLOEQINHEEPTKEMKDKIVAEMDTII 79
 DB 605 ENEKNLNLEKQICMDQAKLEIQKKLSL-----EKSFNQMKNIINEKLTNKEKIK 659
 QY 80 AMIDGVRGVLDRLMQRKDLDFEYQINLEMLKSGDILERDLKKEEARVKKIEV 132
 DB 660 LSLE--IGKLENRINKTAHDKFQSFNYVEKSDMKREKIRKDKIMEKIKM 710
 RESULT 14
 QY17V5 PRELIMINARY; PRT; 311 AA.
 ID QY17V5
 AC QY17V5;
 DT 01-DEC-2001 (TREMELrel. 19, Created)
 DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
 DE Hypothetical protein ST1273.
 GN ST1273.
 OS Sulfolobus tokodaii.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=111955;
 RN [1]
 RQ SEQUENCE FROM N.A.
 RP STRAIN=JCM 10545 / 7;
 RC

```

RX PubMed=11572479;
RA Kavarabasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankaï A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka T., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshihizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermoacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.";
RL DNA Res. 8:123-140(2001).
RW EMBL; AP000985; BAB66315.1; --
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 311 AA; 37536 MW; 674339A5E031B429 CRC64;

Query Match 13.1%; Score 87; DB 17; Length 311;
Best Local Similarity 29.2%; Pred. No. 19;
Matches 28; Conservative 19; Mismatches 41; Indels 8; Gaps 3;

QY 37 RIHQIKKGELALF-----YLQEQNHFEKTKEMDKIVAEEMTIIMIDVGRLDL 92
Db 191 KIYQOLKNERDSLVEKTEILNNQINELVNK--NNEIKNKITEKKDEIKKYRDELKKIDML 249
QY 93 MQRKDLDFEQYNLEMLKKSGDILERDLKKEEARVK 128
Db 250 KSR--NITEVYENQLKNVKNKEIENRKKAEELK 282

RESULT 15
O25166 PRELIMINARY; PRT; 249 AA.
ID AC 025166
AC 025166;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Putative neuraminyllactose-binding hemagglutinin homolog (HPAA).
GN HP0410.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OC NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kervilange A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Loftus B., Richardson J., Zhou L., Kirkness B.F., Peterson S.,
RA Nelson K., Quackenbush J., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori.";
RL Nature 388:539-547(1997).
DR EMBL; AE000557; AAD07478.1; --
RW TIGR; HP0410; --
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 249 AA; 28349 MW; 43E6ABB46EABBS15 CRC64;

Query Match 13.0%; Score 86; DB 16; Length 249;
Best Local Similarity 29.1%; Pred. No. 18;
Matches 34; Conservative 26; Mismatches 39; Indels 18; Gaps 8;

QY 26 YQNEFD---FLMERIHEQTKGELAFYIQEQNHFEKTKEMDK--IVAEMDTIIA 80
Db 77 YQREFEKSFLQSSFLE--RKG-----YSVSQPKDASEIP-QDIKEKALLVLRMDGVA 128
QY 81 MIDVGRLDLRMQRDLDFEQY-NLEMLK-KSGDILER---DLKEEARVKIEV 132
Db 129 ILDEIVESDALGEEKVIDMSSGYLNNFVPEKSDIIHSGFDVSKIKAVIRVEL 185

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2003, 08:17:29 ; Search time 41 Seconds
(without alignments)
511.022 Million cell updates/sec

Title: US-09-877-160A-1

Perfect score: 662

Sequence: 1 MKFTIAFFVATLAVMTSGE.....GDILERDLKKEARVKKIEV 132

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	657	99.2	132	20	AA1980.DAT.*
2	273.5	41.3	134	22	AA1981.DAT.*
3	273.5	41.3	134	23	AA1982.DAT.*
4	273.5	41.3	134	23	AA1983.DAT.*
5	268.5	40.6	134	22	AA1984.DAT.*
6	268.5	40.6	134	23	AA1985.DAT.*
7	262	39.6	127	22	AA1986.DAT.*
8	262	39.6	127	23	AA1987.DAT.*
9	261	39.4	117	22	AA1988.DAT.*

10	261	39.4	117	23	AA19273
11	261	39.4	117	23	AA19275
12	259	39.1	117	22	AAE02248
13	259	39.1	117	22	AAE02249
14	259	39.1	117	22	AAE02253
15	259	39.1	117	23	AAE19271
16	259	39.1	117	23	AAE19276
17	253	38.2	117	22	AAE02250
18	253	38.2	117	22	AAE02252
19	253	38.2	117	23	AAE19272
20	253	38.2	117	23	AAE19274
21	176	26.6	84	22	AAE02258
22	176	26.6	84	23	AAE19281
23	94.5	14.3	506	23	ABP28185
24	94.5	14.3	506	23	ABP28662
25	94	14.2	782	23	AAU98072
26	91	13.7	687	17	AAE97123
27	89	13.4	249	18	AAW55445
28	89	13.4	249	20	AAE17212
29	89	13.4	261	18	AAW55332
30	88	13.3	410	23	ABP43547
31	87.5	13.2	584	22	ABU53081
32	87.5	13.2	623	23	ABP77434
33	87.5	13.2	711	22	ABU53080
34	87.5	13.2	766	22	ABU53068
35	87.5	13.2	907	22	ABU53073
36	87.5	13.2	941	22	ABU53072
37	87.5	13.2	953	22	ABU53069
38	87.5	13.2	2230	24	ABU07445
39	87	13.1	298	23	AAE25662
40	87	13.1	308	23	AAE25655
41	87	13.1	612	22	AAW79138
42	87	13.1	690	22	AAW80122
43	87	13.1	757	23	AAE25653
44	87	13.1	757	23	AAE25671
45	86	13.0	132	23	ABU51422

ALIGNMENTS

RESULT 1
AA1925584

ID AAY25584 standard; protein; 132 AA.

XX AAY25584;

AC AAY25584;

DT 30-SEP-1999 (first entry)

XX D. pteronyssinus allergen Der p 5 protein fragment.

XX Major histocompatibility complex; class II; desensitising; human;
XX allergen; grass; tree; weed; pollen; mould; food; insect; sting;
XX chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
XX screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
XX cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
XX mice; gerbil; vaccine; treatment; prevention; hypersensitivity.

OS Dermatophagoides pteronyssinus.

XX WO9934826-A1.

PD 15-JUL-1999.

XX 11-JAN-1999; 99WO-GB00080.

XX 21-SEP-1998; 98GB-0020474.

XX 09-JAN-1998; 98GB-0000445.

XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.

XX Kay AB, Larche M;

DR WPI; 1999-458255/38.
 XX Desensitizing patients to polypeptide allergens
 XX
 XX Example 6; Page 51; 117pp; English.
 XX This invention describes a novel method of desensitizing a patient to a
 CC polypeptide allergen and comprises administering to the patient a peptide
 CC derived from the allergen where restriction to a MHC Class II molecule
 CC possessed by the patient can be demonstrated for the peptide and the
 CC peptide is able to induce a late phase response in an individual who
 CC possesses the MHC Class II molecule. The methods can be used for
 CC desensitizing patients to allergens present in e.g. grass, tree and weed
 CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
 CC the chitomidae (non-biting midges), spiders and mites, housefly, fruit
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
 CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
 CC produce immunological vaccines which may be used to prevent and/or treat
 CC conditions involving hypersensitivity to allergens. This sequence
 CC represents the house dust mite (Dermatophagoides pteronyssinus) allergen
 CC Der p 5.
 XX
 XX Sequence 132 AA;
 SQ

Query Match 99.2%; Score 657; DB 20; Length 132;
 Best Local Similarity 99.2%; Pred. No. 7.7e-59;
 Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKPFIATVATVAVMTVSGEDKKHGYQNEFDLFLMERIHEQIKKGELALFYIQEQINHFE 60
 DB 1 MKPFIATVATVAVMTVSGEDKKHGYQNEFDLFLMERIHEQIKKGELALFYIQEQINHFE 60
 QY 61 EKPTEKMDKIVAEMDTIIMIDGVGVLDRLMQRKDLDFEQYNLEMKKSGDILERDL 120
 DB 61 EKPTEKMDKIVAEMDTIIMIDGVGVLDRLMQRKDLDFEQYNLEMKKSGDILERDL 120
 QY 121 KKEARVKKIEV 132
 DB 121 KKEARVKKIEV 132

RESULT 2
 AAE02255
 ID AAE02255 standard; Protein; 134 AA.
 XX
 AC AAE02255;
 XX
 DT 31-JUL-2001 (first entry)
 XX
 DE Domestic mite Bt5 polymorphic amino acid variant no.7.
 XX
 KW Mite; immunogenic protein; Bt allergen; therapy; atopic dermatitis;
 KW immediate hypersensitivity; systemic anaphylaxis; allergic rhinitis;
 KW asthma; anti-allergic; anti-inflammatory; immunosuppressive.
 XX
 OS Blomia tropicalis.
 XX
 PN WO200130817-A1.
 XX
 PD 03-MAY-2001.
 XX
 PF 10-OCT-2000; 2000WO-AU01227.
 XX
 PR 26-OCT-1999; 99SG-0005313.
 PR 18-JUL-2000; 2000AU-0008842.
 PR 18-JUL-2000; 2000AU-0008844.
 PR 18-JUL-2000; 2000AU-0008845.
 XX
 PA (UYSI-) UNIV SINGAPORE NAT.
 XX
 PA Chua KY, Cheong N, Lee BW;
 PI

XX
 DR WPI; 2001-309609/32.
 XX N-PSDB; AAD06260.
 XX
 PT Novel immunogenic protein derived from house mite, Blomia tropicalis
 PT useful for treating and diagnosing conditions involving induction of
 PT immuneresponse to mite, such as allergic asthma, atopic dermatitis,
 PT rhinitis
 XX
 PS Example 46; Page 219-220; 230pp; English.
 XX The present invention relates to immunogenic proteins, referred as Bt
 CC allergen, is derived from domestic mite, Blomia tropicalis. The specific
 CC Bt allergens of the invention includes Bt11, Bt10, Bt5 and BtA2. The
 CC immunogenic protein is useful for preventing, reducing or ameliorating
 CC Blomia tropicalis hypersensitivity condition such as atopic dermatitis,
 CC immediate hypersensitivity, systemic anaphylaxis, allergic rhinitis or
 CC asthma and for modulating an immune response directed to Bt allergen in
 CC a subject. The Bt allergens are also useful for detecting antibody
 CC directed to all or a part of Bt allergen in a biological sample from a
 CC subject. Antibodies to Bt allergens are also used as therapeutic or
 CC diagnostic agents, to screen Bt immunoassays and as antagonists to
 CC inhibit Bt activity under circumstances where temporary hypersensitivity
 CC inhibition is required. The present sequence is Bt5 allergen
 CC polymorphic amino acid variant.
 XX
 XX Sequence 134 AA;
 SQ

Query Match 41.3%; Score 273.5; DB 22; Length 134;
 Best Local Similarity 42.5%; Pred. No. 5.4e-20;
 Matches 57; Conservative 32; Mismatches 40; Indels 5; Gaps 2;

QY 1 MKPFIATVATVAVMTVSGEDKKHGYQNEFDLFLMERIHEQIKKGELALFYIQEQIN 57
 DB 1 MKPFIATVATVAVMTVSGEDKKHGYQNEFDLFLMERIHEQIKKGELALFYIQEQIN 57
 QY 58 HFEKPTKMDKIVAEMDTIIMIDGVGVLDRLMQRKDLDFEQYNLEMKKSGDILE 117
 DB 58 ELNENKSKELQKILELDVVCAMIEGAQALEREKLTDLNLERFNYEEAQTLSKILL 118
 QY 118 RDLKKEARVKKIE 131
 DB 119 KDLKTEQKVKDIQ 132

RESULT 3
 AAE19279
 ID AAE19279 standard; Protein; 134 AA.
 XX
 AC AAE19279;
 XX
 DT 21-MAY-2002 (first entry)
 XX
 DE Blomia tropicalis Bt5 polymorphic amino acid variant no. (2)7.
 XX
 KW Immunogenic polymorphic Bt5 protein; domestic mite allergen; rhinitis;
 KW hypersensitivity; immune response; allergy; asthma; atopic dermatitis;
 KW pharmaceutical composition; anti-allergic; antiasthmatic; dermatological;
 KW anti-inflammatory; vaccine; gene therapy; mutant; mutein; variant.
 XX
 OS Blomia tropicalis.
 OS Synthetic.
 XX
 PN WO2000206323-A1.
 XX
 PD 24-JAN-2002.
 XX
 PF 10-OCT-2000; 2000WO-AU01228.
 XX
 PR 18-JUL-2000; 2000AU-0008845.
 XX
 PA (UYSI-) UNIV SINGAPORE NAT.
 XX

```

PI Chua KY, Cheong N, Lee BW, Liew LN;
XX WPI; 2002-195803/25.
DR N-PSDB; AAD30707.
XX
XX Novel immunogenic protein isolated from domestic mite allergen, Blomia
PT tropicalis, useful for preventing, reducing and ameliorating a
PT polymorphic Bt5 hypersensitivity -
XX
XX Claim 4; Page 93; 101pp; English.
XX
XX The present invention relates to an isolated immunogenic polymorphic Bt5
CC protein derived from domestic mite allergen, Blomia tropicalis (Bt). Bt5
CC is useful for preventing, reducing and ameliorating a polymorphic Bt5
CC hypersensitivity. Bt5 is used as vaccine and for modulating an immune
CC response. Bt5 is useful for detecting antibody directed to all or part of
CC polymorphic Bt5 in a biological sample, from a subject, by contacting the
CC biological sample with Bt5, its derivative, homologue, analogue, mimetic
CC or chemical equivalent, to form an antibody-protein complex and detecting
CC the complex. Bt5 is useful for the prophylactic treatment of an allergic
CC condition (including allergic asthma, atopic dermatitis and/or rhinitis)
CC in an individual, e.g. human or animal, by administering Bt5 (preferably
CC in the form of an eukaryotic expression plasmid vector administered in a
CC pharmaceutical composition comprising a normal saline or liposome as
CC carrier), its derivative or homolog, where the airway hyper-reactivity or
CC airway inflammation is prevented. Bt5 gene is useful in gene therapy. The
CC present sequence is Blomia tropicalis Bt5 polymorphic amino acid variant.
XX
XX Sequence 134 AA;
SQ
Query Match 41.3%; Score 273.5; DB 23; Length 134;
Best Local Similarity 42.5%; Pred. No. 5.4e-20;
Matches 57; Conservative 32; Mismatches 40; Indels 5; Gaps 2;
QY 1 MKF---IIAFFVATLAVMTVSGEDKKHDYQNEFDLLMERIHQIKKGELALFYLOEQIN 57
Db 1 MKFAIVLLIACFAA--SVLAQEHKPKDDFRNFDHLLIEQANHAIEKGEHQLLYLOHQLD 58
QY 58 HFEKPTKEMKDKIVAEMDTIIMIDGVGRVLDRLMQRKDLDFEQYNLEMLKKSGLDILE 117
Db 59 ELNENKSKLEQKIIIELDVVCAMIEGAQALEREKRLTDNLILFRNFYEEAQTLSKILL 118
QY 118 RDLKKEEARVKKIE 131
Db 119 KDLKETEQRVKDQIQ 132
RESULT 4
AAE19386
ID AAE19386 standard; Protein; 134 AA.
XX
XX AAE19386;
XX
XX 21-MAY-2002 (first entry)
XX
XX Blomia tropicalis Bt5 amino acid sequence.
XX
XX Immunogenic polymorphic Bt5 protein; domestic mite allergen; rhinitis;
KW hypersensitivity; immune response; allergy; asthma; atopic dermatitis;
KW pharmaceutical composition; antiallergic; antiasthmatic; dermatological;
KW antiinflammatory; vaccine; gene therapy.
XX
XX Blomia tropicalis.
XX
XX WO200206323-A1.
XX
XX 24-JAN-2002.
XX
XX 10-OCT-2000; 2000WO-AU01228.
XX
XX PF 10-OCT-2000; 2000WO-AU01228.
XX
XX PR 18-JUL-2000; 2000AU-000845.
XX
XX PA (UYSI-) UNIV SINGAPORE NAT.
Chua KY, Cheong N, Lee BW, Liew LN;
XX WPI; 2002-195803/25.
DR WPI; 2002-195803/25.
XX
XX Novel immunogenic protein isolated from domestic mite allergen, Blomia
PT tropicalis, useful for preventing, reducing and ameliorating a
PT polymorphic Bt5 hypersensitivity -
XX
XX Disclosure; Fig 3; 101pp; English.
XX
XX The present invention relates to an isolated immunogenic polymorphic Bt5
CC protein derived from domestic mite allergen, Blomia tropicalis (Bt). Bt5
CC is useful for preventing, reducing and ameliorating a polymorphic Bt5
CC hypersensitivity. Bt5 is used as vaccine and for modulating an immune
CC response. Bt5 is useful for detecting antibody directed to all or part of
CC polymorphic Bt5 in a biological sample, from a subject, by contacting the
CC biological sample with Bt5, its derivative, homologue, analogue, mimetic
CC or chemical equivalent, to form an antibody-protein complex and detecting
CC the complex. Bt5 is useful for the prophylactic treatment of an allergic
CC condition (including allergic asthma, atopic dermatitis and/or rhinitis)
CC in an individual, e.g. human or animal, by administering Bt5 (preferably
CC in the form of an eukaryotic expression plasmid vector administered in a
CC pharmaceutical composition comprising a normal saline or liposome as
CC carrier), its derivative or homolog, where the airway hyper-reactivity or
CC airway inflammation is prevented. Bt5 gene is useful in gene therapy. The
CC present sequence is Blomia tropicalis Bt5 amino acid sequence.
XX
XX Sequence 134 AA;
SQ
Query Match 41.3%; Score 273.5; DB 23; Length 134;
Best Local Similarity 42.5%; Pred. No. 5.4e-20;
Matches 57; Conservative 32; Mismatches 40; Indels 5; Gaps 2;
QY 1 MKF---IIAFFVATLAVMTVSGEDKKHDYQNEFDLLMERIHQIKKGELALFYLOEQIN 57
Db 1 MKFAIVLLIACFAA--SVLAQEHKPKDDFRNFDHLLIEQANHAIEKGEHQLLYLOHQLD 58
QY 58 HFEKPTKEMKDKIVAEMDTIIMIDGVGRVLDRLMQRKDLDFEQYNLEMLKKSGLDILE 117
Db 59 ELNENKSKLEQKIIIELDVVCAMIEGAQALEREKRLTDNLILFRNFYEEAQTLSKILL 118
QY 118 RDLKKEEARVKKIE 131
Db 119 KDLKETEQRVKDQIQ 132
RESULT 5
AAE02256
ID AAE02256 standard; Protein; 134 AA.
XX
XX AAE02256;
XX
XX 31-JUL-2001 (first entry)
XX
XX Domestic mite Bt5 polymorphic amino acid variant no.24.
XX
XX Mite; immunogenic protein; Bt allergen; therapy; atopic dermatitis;
KW immediate hypersensitivity; systemic anaphylaxis; allergic rhinitis;
KW asthma; antiallergic; antiinflammatory; immunosuppressive.
XX
XX Blomia tropicalis.
XX
XX WO200130817-A1.
XX
XX 03-MAY-2001.
XX
XX 10-OCT-2000; 2000WO-AU01227.
XX
XX PF 10-OCT-2000; 2000WO-AU01227.
XX
XX PR 26-OCT-1999; 99SG-0005313.
XX
XX PR 18-JUL-2000; 2000AU-000842.
XX
XX PR 18-JUL-2000; 2000AU-000844.
XX
XX PR 18-JUL-2000; 2000AU-000845.

```

```

XX (UYSI-) UNIV SINGAPORE NAT.
XX Chua KY, Cheong N, Lee BW;
XX WPI; 2001-308609/32.
XX N-PSDB; AAD06267.
XX Novel immunogenic protein derived from house mite, Blomia tropicalis
XX useful for treating and diagnosing conditions involving induction of
XX immuneresponse to mite, such as allergic asthma, atopic dermatitis,
XX rhinitis
XX Example 46; Page 220-221; 230pp; English.
XX The present invention relates to immunogenic proteins, referred as Bt
XX allergen, is derived from domestic mite, Blomia tropicalis. The specific
XX Bt allergens of the invention includes Bt11, Bt10, Bt5 and BtA2. The
XX immunogenic protein is useful for preventing, reducing or ameliorating
XX Blomia tropicalis hypersensitivity condition such as atopic dermatitis,
XX immediate hypersensitivity, systemic anaphylaxis, allergic rhinitis or
XX asthma and for modulating an immune response directed to Bt allergen in
XX a subject. The Bt allergens are also useful for detecting antibody
XX directed to all or a part of Bt allergen in a biological sample from a
XX subject. Antibodies to Bt allergens are also used as therapeutic or
XX diagnostic agents, to screen Bt immunoassays and as antagonists to
XX inhibit Bt activity under circumstances where temporary hypersensitivity
XX inhibition is required. The present sequence is Bt5 allergen
XX polymorphic amino acid variant.
XX Sequence 134 AA;
SQ
Query Match 40.6%; Score 268.5; DB 22; Length 134;
Best Local Similarity 41.8%; Pred. No. 1.7e-19;
Matches 56; Conservative 32; Mismatches 41; Indels 5; Gaps 2;
QY 1 MKF---IIAFVATLAVMTVSGEDKKHYQNEFDPLMERIHEQIKKGELALFYLOEQIN 57
DB 1 MKFAIVLIACFAA--SVLAQEHKPKDDFRNEFDHLLIEQANHAIENGHEQLLYLQHOLD 58
QY 58 HFEKPTKEMDKIVAEMDTIIAMIDVGVLDRLMQRKDLDFEQYNLEMLKSGDILE 117
DB 59 ELNENKSELQEKIIRLEDVVCAMIEGAQALERELKRTDLNILERFNVEEAQTLSKILL 118
QY 118 RDLKKEEARVKKIE 131
DB 119 KDLKETEOKVKDIQ 132
RESULT 6
AAE19280
ID AAE19280 standard; Protein; 134 AA.
AC AAE19280;
XX
DT 21-MAY-2002 (first entry)
XX
DE Blomia tropicalis Bt5 polymorphic amino acid variant no. (9)24.
XX
KW Immunogenic polymorphic Bt5 protein; domestic mite allergen; rhinitis;
KW hypersensitivity; immune response; allergy; asthma; atopic dermatitis;
KW pharmaceutical composition; antiallergic; antiasthmatic; dermatological;
KW antiinflammatory; vaccine; gene therapy; mutant; mutein; variant.
XX
OS Blomia tropicalis.
OS Synthetic.
XX
PN WO200206323-A1.
XX
PD 24-JAN-2002.
XX
PF 10-OCT-2000; 2000WO-AU01228.
XX
XX

```

```

PR 18-JUL-2000; 2000AU-0008845.
XX (UYSI-) UNIV SINGAPORE NAT.
XX Chua KY, Cheong N, Lee BW, Liew LN;
XX WPI; 2002-195803/25.
XX N-PSDB; AAD30714.
XX Novel immunogenic protein isolated from domestic mite allergen, Blomia
XX tropicalis, useful for preventing, reducing and ameliorating a
XX polymorphic Bt5 hypersensitivity
XX Claim 4; Page 93-94; 101pp; English.
XX The present invention relates to an isolated immunogenic polymorphic Bt5
XX protein derived from domestic mite allergen, Blomia tropicalis (Bt). Bt5
XX is useful for preventing, reducing and ameliorating a polymorphic Bt5
XX hypersensitivity. Bt5 is used as vaccine and for modulating an immune
XX response. Bt5 is useful for detecting antibody directed to all or part of
XX polymorphic Bt5 in a biological sample, from a subject, by contacting the
XX biological sample with Bt5, its derivative, homologue, analogue, mimetic
XX or chemical equivalent, to form an antibody-protein complex and detecting
XX the complex. Bt5 is useful for the prophylactic treatment of an allergic
XX condition (including allergic asthma, atopic dermatitis and/or rhinitis)
XX in an individual, e.g. human or animal, by administering Bt5 (preferably
XX in the form of an eukaryotic expression plasmid vector administered in a
XX pharmaceutical composition comprising a normal saline or liposome as
XX carrier), its derivative or homolog, where the airway hyper-reactivity or
XX airway inflammation is prevented. Bt5 gene is useful in gene therapy. The
XX present sequence is Blomia tropicalis Bt5 polymorphic amino acid variant.
XX Sequence 134 AA;
SQ
Query Match 40.6%; Score 268.5; DB 23; Length 134;
Best Local Similarity 41.8%; Pred. No. 1.7e-19;
Matches 56; Conservative 32; Mismatches 41; Indels 5; Gaps 2;
QY 1 MKF---IIAFVATLAVMTVSGEDKKHYQNEFDPLMERIHEQIKKGELALFYLOEQIN 57
DB 1 MKFAIVLIACFAA--SVLAQEHKPKDDFRNEFDHLLIEQANHAIENGHEQLLYLQHOLD 58
QY 58 HFEKPTKEMDKIVAEMDTIIAMIDVGVLDRLMQRKDLDFEQYNLEMLKSGDILE 117
DB 59 ELNENKSELQEKIIRLEDVVCAMIEGAQALERELKRTDLNILERFNVEEAQTLSKILL 118
QY 118 RDLKKEEARVKKIE 131
DB 119 KDLKETEOKVKDIQ 132
RESULT 7
AAE02257
ID AAE02257 standard; Protein; 127 AA.
XX
AC AAE02257;
XX
DT 31-JUL-2001 (first entry)
XX
DE Domestic mite Bt5 polymorphic amino acid variant no.13.
XX
KW Mite; immunogenic protein; Bt allergen; therapy; atopic dermatitis;
KW immediate hypersensitivity; systemic anaphylaxis; allergic rhinitis;
KW asthma; antiallergic; antiinflammatory; immunosuppressive.
XX
OS Blomia tropicalis.
XX
PN WO200130817-A1.
XX
PD 03-MAY-2001.
XX
PF 10-OCT-2000; 2000WO-AU01227.
XX
XX

```



```

PF 10-OCT-2000; 2000WO-AU01227.
XX
PR 26-OCT-1999; 99SG-0005313.
PR 18-JUL-2000; 2000AU-0008842.
PR 18-JUL-2000; 2000AU-0008844.
PR 18-JUL-2000; 2000AU-0008845.
XX
PA (UYSI-) UNIV SINGAPORE NAT.
XX
PI Chua KY, Cheong N, Lee BW;
XX WPI; 2001-308609/32.
DR N-PSDB; AAD06252.
XX
PT Novel immunogenic protein derived from house mite, Blomia tropicalis
PT useful for treating and diagnosing conditions involving induction of
PT immuneresponse to mite, such as allergic asthma, atopic dermatitis,
PT rhinitis
XX
PS Example 46; Page 210-211; 230pp; English.
XX
CC The present invention relates to immunogenic proteins, referred as Bt
CC allergen, is derived from domestic mite, Blomia tropicalis. The specific
CC Bt allergens of the invention includes Bt11, Bt10, Bt5 and BtA2. The
CC immunogenic protein is useful for preventing, reducing or ameliorating
CC Blomia tropicalis hypersensitivity condition such as atopic dermatitis,
CC immediate hypersensitivity, systemic anaphylaxis, allergic rhinitis or
CC asthma and for modulating an immune response directed to Bt allergen in
CC a subject. The Bt allergens are also useful for detecting antibody
CC directed to all or a part of Bt allergen in a biological sample from a
CC subject. Antibodies to Bt allergens are also used as therapeutic or
CC diagnostic agents, to screen Bt immunoassays and as antagonists to
CC inhibit Bt activity under circumstances where temporary hypersensitivity
CC inhibition is required. The present sequence is mite Bt5 polymorphic
CC protein.
XX
SQ Sequence 117 AA;
Query Match 39.4%; Score 261; DB 22; Length 117;
Best Local Similarity 43.8%; Pred. No. 8.5e-19;
Matches 49; Conservative 29; Mismatches 34; Indels 0; Gaps 0;
QY 20 EDKXHDYQNEFDLLMERIHEQIKKGELALFYIQEINHFEEKPTKEMKDKIVAEMDTII 79
DB 4 EPKDDFRNEFDHLLIEQANHAIEKGEHOILYQHQDELNENKSKELQEKIREDLVVC 63
QY 80 AMIDGVGVLDRLMORKDLDFEQYNLEMLKXSGDILERDLKKEARVKKIE 131
DB 64 AMIEGAQGALERELKRTDLNILERFNVEEAQTLISKILLKDLKETEOKVKDQ 115
RESULT 10
AAE19273
ID AAE19273 standard; Protein; 117 AA.
XX
AC AAE19273;
XX
DT 21-MAY-2002 (first entry)
XX
DE Blomia tropicalis Bt5 polymorphic amino acid sequence no.5.
XX
KW Immunogenic polymorphic Bt5 protein; domestic mite allergen; rhinitis;
KW hypersensitivity; immune response; allergy; asthma; atopic dermatitis;
KW pharmaceutical composition; antiallergic; antiasthmatic; dermatological;
KW antiinflammatory; vaccine; gene therapy.
XX
OS Blomia tropicalis.
XX
FH Key Location/Qualifiers
FT Misc-difference 21..22
FT /note= "Encoded by CAGTAAACAGTITTTTTTGAAGGT
FT TAAATAAAAGTAATCAAGAACTTTTAGGCA"
FT Misc-difference 33

```

```

FT FT /note= "Encoded by TTG"
FT Misc-difference 34
XX /note= "Encoded by ATT"
XX WO200206323-A1.
XX 24-JAN-2002.
XX
PF 10-OCT-2000; 2000WO-AU01228.
XX
PR 18-JUL-2000; 2000AU-0008845.
XX
PA (UYSI-) UNIV SINGAPORE NAT.
XX
PI Chua KY, Cheong N, Lee BW, Liew LN;
XX WPI; 2002-195803/25.
DR N-PSDB; AAD30699.
XX
PT Novel immunogenic protein isolated from domestic mite allergen, Blomia
PT tropicalis, useful for preventing, reducing and ameliorating a
PT polymorphic Bt5 hypersensitivity
XX
PS Claim 4; Page 84-85; 101pp; English.
XX
CC The present invention relates to an isolated immunogenic polymorphic Bt5
CC protein derived from domestic mite allergen, Blomia tropicalis (Bt). Bt5
CC is useful for preventing, reducing and ameliorating a polymorphic Bt5
CC hypersensitivity. Bt5 is used as vaccine and for modulating an immune
CC response. Bt5 is useful for detecting antibody directed to all or part of
CC biological Bt5 in a biological sample, from a subject, by contacting the
CC biological sample with Bt5, its derivative, from a subject, analogue, mimetic
CC or chemical equivalent, to form an antibody-protein complex and detecting
CC the complex. Bt5 is useful for the prophylactic treatment of an allergic
CC condition (including allergic asthma, atopic dermatitis and/or rhinitis)
CC in an individual, e.g. human or animal, by administering Bt5 (preferably
CC in the form of an eukaryotic expression plasmid vector administered in a
CC pharmaceutical composition comprising a normal saline or liposome as
CC carrier), its derivative or homolog, where the airway hyper-reactivity
CC or airway inflammation is prevented. Bt5 gene is useful in gene therapy.
CC The present sequence is Blomia tropicalis Bt5 polymorphic amino acid
CC sequence.
XX
SQ Sequence 117 AA;
Query Match 39.4%; Score 261; DB 23; Length 117;
Best Local Similarity 43.8%; Pred. No. 8.5e-19;
Matches 49; Conservative 29; Mismatches 34; Indels 0; Gaps 0;
QY 20 EDKXHDYQNEFDLLMERIHEQIKKGELALFYIQEINHFEEKPTKEMKDKIVAEMDTII 79
DB 4 EPKDDFRNEFDHLLIEQANHAIEKGEHOILYQHQDELNENKSKELQEKIREDLVVC 63
QY 80 AMIDGVGVLDRLMORKDLDFEQYNLEMLKXSGDILERDLKKEARVKKIE 131
DB 64 AMIEGAQGALERELKRTDLNILERFNVEEAQTLISKILLKDLKETEOKVKDQ 115
RESULT 11
AAE19275
ID AAE19275 standard; Protein; 117 AA.
XX
AC AAE19275;
XX
DT 21-MAY-2002 (first entry)
XX
DE Blomia tropicalis Bt5 polymorphic amino acid sequence no.11.
XX
KW Immunogenic polymorphic Bt5 protein; domestic mite allergen; rhinitis;
KW hypersensitivity; immune response; allergy; asthma; atopic dermatitis;
KW pharmaceutical composition; antiallergic; antiasthmatic; dermatological;
KW antiinflammatory; vaccine; gene therapy.
XX
OS Blomia tropicalis.
XX
FH Key Location/Qualifiers
FT Misc-difference 21..22
FT /note= "Encoded by CAGTAAACAGTITTTTTTGAAGGT
FT TAAATAAAAGTAATCAAGAACTTTTAGGCA"
FT Misc-difference 33

```


OS Blomia tropicalis.
 XX Key Location/Qualifiers
 FT Misc-difference 4
 FT Misc-difference /note= "Encoded by AG"
 FT Misc-difference 23..24
 FT /note= "Encoded by AACAGTTTTTTTGAAGGTTAATAAAAAAGTAAT
 FT CAAAGAACTTTTGTAGCAACCAT"
 XX
 PN WO200206323-A1.
 XX
 XX
 PD 24-JAN-2002.
 XX
 XX 10-OCT-2000; 2000WO-AU01228.
 XX
 XX 18-JUL-2000; 2000AU-0008845.
 XX
 XX (UYSI-) UNIV SINGAPORE NAT.
 PA
 XX Chua KY, Cheong N, Lee BW, Liew LN;
 PI
 XX WPI: 2002-195803/25.
 DR
 DR N-PSDB; AAD30705.
 XX
 XX Novel immunogenic protein isolated from domestic mite allergen, Blomia
 PT tropicalis, useful for preventing, reducing and ameliorating a
 PT polymorphic Bt5 hypersensitivity -
 XX
 XX Claim 4; Page 86; 101pp; English.
 PS
 XX The present invention relates to an isolated immunogenic polypeptide Bt5
 CC protein derived from domestic mite allergen, Blomia tropicalis (Bt). Bt5
 CC is useful for preventing, reducing and ameliorating a polymorphic Bt5
 CC hypersensitivity. Bt5 is used as vaccine and for modulating an immune
 CC response. Bt5 is useful for detecting antibody directed to all or part of
 CC polymorphic Bt5 in a biological sample, from a subject, by contacting the
 CC biological sample with Bt5, its derivative, homologue, analogue, mimetic
 CC or chemical equivalent, to form an antibody-protein complex and detecting
 CC the complex. Bt5 is useful for the prophylactic treatment of an allergic
 CC condition (including allergic asthma, atopic dermatitis and/or rhinitis)
 CC in an individual, e.g. human or animal, by administering Bt5 (preferably
 CC in the form of an eukaryotic expression plasmid vector administered in a
 CC pharmaceutical composition comprising a normal saline or liposome as
 CC carrier), its derivative or homolog, where the airway hyper-reactivity or
 CC airway inflammation is prevented. Bt5 gene is useful in gene therapy. The
 CC present sequence is Blomia tropicalis Bt5 polymorphic amino acid
 CC sequence.
 XX
 SQ Sequence 117 AA;
 Query Match 39.4%; Score 261; DB 23; Length 117;
 Best Local Similarity 43.8%; Pred. No. 8.5e-19;
 Matches 49; Conservative 29; Mismatches 34; Indels 0; Gaps 0;
 Qy 20 EDKHYQNEPFLMERIHQIKKGELALFYLOEQINHFPEKPTKEMKDKIVAEMDTII 79
 Db 4 EPKXDFRNEPDLHIEQANHAIEKGEHQILYLOHLDENLNKSKELQEKIIRLDVVC 63
 Qy 80 AMIDGVGLDRLMQKDLDFEQVNLMLKSGDILERDLKKEARVKIE 131
 Db 64 AMIEGAQALRELRKTDNLILRFNYEEAQTLSKILKDLKETEOKVKDIQ 115
 RESULT 12
 AAE02248
 ID AAE02248 standard; Protein; 117 AA.
 AC
 AC AAE02248;
 XX
 XX 31-JUL-2001 (first entry)
 DT
 XX Domestic mite Bt5 allergen.
 DE
 XX

KW Mite; immunogenic protein; Bt allergen; therapy; atopic dermatitis;
 KW immediate hypersensitivity; systemic anaphylaxis; allergic rhinitis;
 KW asthma; anti-allergic; anti-inflammatory; immunosuppressive.
 XX
 OS Blomia tropicalis.
 XX
 PN WO200130817-A1.
 XX
 PD 03-MAY-2001.
 XX
 XX 10-OCT-2000; 2000WO-AU01227.
 XX
 XX 26-OCT-1999; 99SG-0005313.
 PR
 PR 18-JUL-2000; 2000AU-0008842.
 PR
 PR 18-JUL-2000; 2000AU-0008844.
 PR
 PR 18-JUL-2000; 2000AU-0008845.
 XX
 XX (UYSI-) UNIV SINGAPORE NAT.
 PA
 XX Chua KY, Cheong N, Lee BW;
 PI
 XX WPI: 2001-308609/32.
 DR
 DR N-PSDB; AAD06247.
 XX
 XX Novel immunogenic protein derived from house mite, Blomia tropicalis
 PT useful for treating and diagnosing conditions involving induction of
 PT immuneresponse to mite, such as allergic asthma, atopic dermatitis,
 PT rhinitis -
 XX
 XX Example 45; Fig 20; 230pp; English.
 PS
 XX The present invention relates to immunogenic proteins, referred as Bt
 CC allergen, is derived from domestic mite, Blomia tropicalis. The specific
 CC Bt allergens of the invention includes Bt11, Bt10, Bt5 and BtA2. The
 CC immunogenic protein is useful for preventing, reducing or ameliorating
 CC Blomia tropicalis hypersensitivity condition such as atopic dermatitis or
 CC immediate hypersensitivity, systemic anaphylaxis, allergic rhinitis or
 CC asthma and for modulating an immune response directed to Bt allergen in
 CC a subject. The Bt allergens are also useful for detecting antibody
 CC directed to all or a part of Bt allergen in a biological sample from a
 CC subject. Antibodies to Bt allergens are also used as therapeutic or
 CC diagnostic agents, to screen Bt immunoassays and as antagonists to
 CC inhibit Bt activity under circumstances where temporary hypersensitivity
 CC inhibition is required. The present sequence is Bt5 allergen.
 XX
 SQ Sequence 117 AA;
 Query Match 39.1%; Score 259; DB 22; Length 117;
 Best Local Similarity 44.5%; Pred. No. 1.4e-18;
 Matches 49; Conservative 28; Mismatches 33; Indels 0; Gaps 0;
 Qy 22 KKHQYQNEPFLMERIHQIKKGELALFYLOEQINHFPEKPTKEMKDKIVAEMDTIIAM 81
 Db 6 KDDFRNEPDLHIEQANHAIEKGEHQILYLOHLDENLNKSKELQEKIIRLDVVCAM 65
 Qy 82 IDGVGVGLDRLMQKDLDFEQVNLMLKSGDILERDLKKEARVKIE 131
 Db 66 IEQAQALRELRKTDNLILRFNYEEAQTLSKILKDLKETEOKVKDIQ 115
 RESULT 13
 AAE02249
 ID AAE02249 standard; Protein; 117 AA.
 XX
 AC AAE02249;
 XX
 XX 31-JUL-2001 (first entry)
 DT
 XX Domestic mite Bt5 polymorphic amino acid sequence no: 1-3 and 7-10.
 DE
 XX Mite; immunogenic protein; Bt allergen; therapy; atopic dermatitis;
 KW immediate hypersensitivity; systemic anaphylaxis; allergic rhinitis;
 KW asthma; anti-allergic; anti-inflammatory; immunosuppressive.

```

XX OS Blomia tropicalis.
XX FH Key Location/Qualifiers
XX FT Misc-difference 17 /note= "Encoded by TGG"
XX FT Misc-difference 34 /note= "Encoded by ATT"
XX FT Misc-difference 35 /note= "Encoded by CAC"
XX PN WO200130817-A1.
XX PD 03-MAY-2001.
XX PF 10-OCT-2000; 2000WO-AU01227.
XX PR 26-OCT-1999; 99SG-0005313.
XX PR 18-JUL-2000; 2000AU-0008842.
XX PR 18-JUL-2000; 2000AU-0008844.
XX PR 18-JUL-2000; 2000AU-0008845.
XX PA (UYSI-) UNIV SINGAPORE NAT.
XX PI Chua XY, Cheong N, Lee BW;
XX DR N-PSDB; AAD06248, AAD06249, AAD06250, AAD06254, AAD06256, AAD06257.
XX DR WPI; 2001-308609/32.
XX PT Novel immunogenic protein derived from house mite, Blomia tropicalis
XX PT useful for treating and diagnosing conditions involving induction of
XX PT immuneresponse to mite, such as allergic asthma, atopic dermatitis,
XX PT rhinitis -
XX PS Example 46; Page 209; 230pp; English.
XX CC The present invention relates to immunogenic proteins, referred as Bt
XX CC allergen, is derived from domestic mite, Blomia tropicalis. The specific
XX CC Bt allergens of the invention includes Bt11, Bt10, Bt5 and BtA2. The
XX CC immunogenic protein is useful for preventing, reducing or ameliorating
XX CC Blomia tropicalis hypersensitivity condition such as atopic dermatitis,
XX CC immediate hypersensitivity, systemic anaphylaxis, allergic rhinitis or
XX CC asthma and for modulating an immune response directed to Bt allergen in
XX CC a subject. The Bt allergens are also useful for detecting antibody
XX CC directed to all or a part of Bt allergen in a biological sample from a
XX CC subject. Antibodies to Bt allergens are also used as therapeutic or
XX CC diagnostic agents, to screen Bt immunoassays and as antagonists to
XX CC inhibit Bt activity under circumstances where temporary hypersensitivity
XX CC inhibition is required. The present sequence is mite Bt5 polymorphic
XX CC protein.
XX CC Sequence 117 AA;
XX CC Query Match 39.1%; Score 259; DB 22; Length 117;
XX CC Best Local Similarity 44.5%; Pred. No. 1.4e-18;
XX CC Matches 49; Conservative 28; Mismatches 33; Indels 0; Gaps 0;
XX QY 22 KKHVDYNEFDLMEIRHEQIKKGELALFYLQEQINHFEEKPTKEMKDKIVAEEMDTIAM 81
XX DB 6 KKDDFRNEFDHLLEIQANHAIEKGEHQLLYQHQDLNENKSKELQEKIRIREDLVVVCAM 65
XX QY 82 IDGVRGVLDRMQKDLDFEQYNLEMLKKGSDILERDLKKEARVKKIE 131
XX DB 66 IEGAQALEREKRLKTDNLIERNFVVEEAQTLISKILLKDLKETEOKVKDQ 115
XX RESULT 14
XX ID AAE02253
XX ID AAE02253 standard; Protein; 117 AA.
XX AC AAE02253;
XX DT 31-JUL-2001 (first entry)

```

```

XX DE Domestic Bt5 polymorphic amino acid sequence no: 11.
XX KW Mite; immunogenic protein; Bt allergen; therapy; atopic dermatitis;
XX KW immediate hypersensitivity; systemic anaphylaxis; allergic rhinitis;
XX KW asthma; antiallergic; antiinflammatory; immunosuppressive.
XX OS Blomia tropicalis.
XX PN WO200130817-A1.
XX PD 03-MAY-2001.
XX PF 10-OCT-2000; 2000WO-AU01227.
XX PR 26-OCT-1999; 99SG-0005313.
XX PR 18-JUL-2000; 2000AU-0008842.
XX PR 18-JUL-2000; 2000AU-0008844.
XX PR 18-JUL-2000; 2000AU-0008845.
XX PA (UYSI-) UNIV SINGAPORE NAT.
XX PI Chua XY, Cheong N, Lee BW;
XX DR N-PSDB; AAD06258.
XX PT Novel immunogenic protein derived from house mite, Blomia tropicalis
XX PT useful for treating and diagnosing conditions involving induction of
XX PT immuneresponse to mite, such as allergic asthma, atopic dermatitis,
XX PT rhinitis -
XX PS Example 46; Page 212; 230pp; English.
XX CC The present invention relates to immunogenic proteins, referred as Bt
XX CC allergen, is derived from domestic mite, Blomia tropicalis. The specific
XX CC Bt allergens of the invention includes Bt11, Bt10, Bt5 and BtA2. The
XX CC immunogenic protein is useful for preventing, reducing or ameliorating
XX CC Blomia tropicalis hypersensitivity condition such as atopic dermatitis,
XX CC immediate hypersensitivity, systemic anaphylaxis, allergic rhinitis or
XX CC asthma and for modulating an immune response directed to Bt allergen in
XX CC a subject. The Bt allergens are also useful for detecting antibody
XX CC directed to all or a part of Bt allergen in a biological sample from a
XX CC subject. Antibodies to Bt allergens are also used as therapeutic or
XX CC diagnostic agents, to screen Bt immunoassays and as antagonists to
XX CC inhibit Bt activity under circumstances where temporary hypersensitivity
XX CC inhibition is required. The present sequence is mite Bt5 polymorphic
XX CC protein.
XX CC Sequence 117 AA;
XX CC Query Match 39.1%; Score 259; DB 22; Length 117;
XX CC Best Local Similarity 44.5%; Pred. No. 1.4e-18;
XX CC Matches 49; Conservative 28; Mismatches 33; Indels 0; Gaps 0;
XX QY 22 KKHVDYNEFDLMEIRHEQIKKGELALFYLQEQINHFEEKPTKEMKDKIVAEEMDTIAM 81
XX DB 6 KKDDFRNEFDHLLEIQANHAIEKGEHQLLYQHQDLNENKSKELQEKIRIREDLVVVCAM 65
XX QY 82 IDGVRGVLDRMQKDLDFEQYNLEMLKKGSDILERDLKKEARVKKIE 131
XX DB 66 IEGAQALEREKRLKTDNLIERNFVVEEAQTLISKILLKDLKETEOKVKDQ 115
XX RESULT 15
XX ID AAE19271
XX ID AAE19271 standard; Protein; 117 AA.
XX AC AAE19271;
XX DT 21-MAY-2002 (first entry)
XX DT Blomia tropicalis Bt5 amino acid sequence (ref 12).

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2003, 08:24:35 ; Search time 29 Seconds
(without alignments)
830.960 Million cell updates/sec

Title: US-09-877-160A-1

Perfect score: 662

Sequence: 1 MKFIIAFFVATLAVMTVSGE.....GDLERDLKKEARVKKIEV 132

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 666188 seqs, 18259486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	657	99.2	132	10	US-09-877-160-1
2	657	99.2	132	11	US-09-847-208-83
3	86	13.0	249	10	US-09-881-752A-264
4	86	13.0	249	12	US-10-080-113-1
5	82	12.4	379	12	US-10-100-699-2
6	82	12.4	380	9	US-09-134-333-13
7	82	12.4	1786	10	US-09-742-096-3
8	80.5	12.2	676	12	US-10-256-250-16
9	80.5	12.2	709	12	US-10-256-250-15
10	80.5	12.2	725	11	US-09-978-309A-47
11	80.5	12.2	725	12	US-10-256-250-14
12	80.5	12.2	2383	15	US-10-082-830-260
13	80	12.1	1938	15	US-10-171-311-164
14	80	12.1	1945	11	US-09-927-597-2
15	80	12.1	1972	15	US-10-171-311-162

16	80	12.1	1979	11	US-09-927-597-4
17	78	11.8	862	15	US-10-177-293-391
18	77.5	11.7	475	10	US-09-738-973-193
19	77.5	11.7	475	10	US-09-854-133-193
20	77.5	11.7	475	15	US-10-144-649A-193
21	77.5	11.7	1055	9	US-09-815-242-11217
22	77	11.6	380	9	US-09-134-333-12
23	76.5	11.6	107	15	US-10-023-066A-92
24	76.5	11.6	419	11	US-09-863-049A-2
25	76.5	11.6	419	12	US-09-851-673-2
26	76	11.6	431	12	US-10-254-995-3
27	75.5	11.4	234	11	US-09-819-142-16
28	75.5	11.4	234	11	US-09-934-455-358
29	75.5	11.4	967	15	US-10-197-666A-90
30	75.5	11.4	1591	15	US-10-197-666A-92
31	75	11.3	1707	15	US-10-128-714-3556
32	75	11.3	1750	15	US-10-128-714-8556
33	74.5	11.3	488	12	US-10-032-585-7710
34	74.5	11.3	821	10	US-09-884-001-18
35	74.5	11.3	975	15	US-10-097-340-31
36	74	11.2	281	12	US-10-080-608A-63
37	74	11.2	307	11	US-09-298-523B-7
38	74	11.2	715	10	US-09-870-759-86
39	74	11.2	715	12	US-09-751-708A-86
40	74	11.2	2125	10	US-09-919-172-29
41	74	11.2	2649	12	US-10-205-219-169
42	73.5	11.1	208	15	US-09-925-637-52
43	73.5	11.1	208	15	US-10-084-205-52
44	73.5	11.1	251	12	US-10-254-995-4
45	73.5	11.1	254	11	US-09-056-019-9

ALIGNMENTS

RESULT 1

US-09-877-160-1
; Sequence 1, Application US/09877160
; Publication No. US20020197268A1
; GENERAL INFORMATION:
; APPLICANT: Cheng-Huang, Hsu
; APPLICANT: Winston T. K.
; TITLE OF INVENTION: ALLERGEN-CONTAINING MILK FOR ALLERGY
; TITLE OF INVENTION: TREATMENT
; FILE REFERENCE: 12774-003001
; CURRENT APPLICATION NUMBER: US/09/877,160
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-09-877-160-1

Query Match 99.2%; Score 657; DB 10; Length 132;
Best Local Similarity 99.2%; Pred. No. 2.9e-59;
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	MKFIIAFFVATLAVMTVS	GDKKHQYQNEFD	FLMERIHEQIKKGELALFYIQEQINHFE	60
Db	1	MKFIIAFFVATLAVMTVS	GDKKHQYQNEFD	FLMERIHEQIKKGELALFYIQEQINHFE	60
Qy	61	EKPTKMDKIVAENDTII	AMIDGVGRVLDRLMQ	RKDLDFEQYNLEMLKKS	GDILERDL 120
Db	61	EKPTKMDKIVAENDTII	AMIDGVGRVLDRLMQ	RKDLDFEQYNLEMLKKS	GDILERDL 120
Qy	121	KKEARVKKIEV			132
Db	121	KKEARVKKIEV			132

RESULT 2

```
US-09-847-208-83
; Sequence 83, Application US/09847208
; Publication No. US20030082190A1
; GENERAL INFORMATION:
; APPLICANT: Saxon, Andrew
; APPLICANT: Zhang, Ke
; APPLICANT: Zhu, Daoheng
; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
; TITLE OF INVENTION: IGE-MEDIATED ALLERGIC DISEASES
; FILE REFERENCE: US67,002A
; CURRENT APPLICATION NUMBER: US/09/847,208
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 83
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Dermatophagoides pteromyssinus (House-dust mite)
US-09-847-208-83

Query Match          99.2%; Score 657; DB 11; Length 132;
Best Local Similarity 99.2%; Pred. No. 2.9e-59;
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKFIIAFFVATLAVMTVSGEDKKHDYQNEFDLIMERIHEQIKKGELALFYLOEQINHFE 60
DB 1 MKFIIAFFVATLAVMTVSGEDKKHDYQNEFDLIMERIHEQIKKGELALFYLOEQINHFE 60
QY 61 EKTETKMKDQIVAMDTIIMIDGVRGVLDRMQKDLDFEQYNLEMLKSGDILERDL 120
DB 61 EKTETKMKDQIVAMDTIIMIDGVRGVLDRMQKDLDFEQYNLEMLKSGDILERDL 120
QY 121 KKEEARVKKIEV 132
DB 121 KKEEARVKKIEV 132

RESULT 3
US-09-881-752A-264
; Sequence 264, Application US/09881752A
; Patent No. US20020115078A1
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Comen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20020115078A1 Helicobacter Polypeptides in the
; FILE REFERENCE: Genome
; CURRENT APPLICATION NUMBER: US/09/881,752A
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 08/833,457
; PRIOR FILING DATE: 1997-04-01
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 264
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-881-752A-264

Query Match          13.0%; Score 86; DB 10; Length 249;
Best Local Similarity 29.1%; Pred. No. 0.68;
Matches 34; Conservative 26; Mismatches 39; Indels 18; Gaps 8;

QY 26 YQNEFD---FLMERIHEQIKKGELALFYLOEQINHEFEKPKTKEMKDK--IVAEMDTIIA 80
DB 77 YQKEFEKSLFLQSSSLE--RKG-----YVSQPKDASEIP-QDIKEKALLVLRMDGNTA 128
QY 81 MIDGVRGVLDRMQKDLDFEQY-NLEMLK-KSGDILER---DLKKEARVKKIEV 132
DB 81 MIDGVRGVLDRMQKDLDFEQY-NLEMLK-KSGDILER---DLKKEARVKKIEV 132

US-09-877-160a-1.rapb
```

```
Db 129 ILEDIVESDALSEEKVIDMSSGYLNLNFVEPKSEDIHHSFGIDVSKIKAVIERVEL 185

RESULT 4
US-10-080-113-1
; Sequence 1, Application US/10080113
; Publication No. US20030166027A1
; GENERAL INFORMATION:
; APPLICANT: SACHS, GEORGE
; APPLICANT: VOLAND, PETRA
; TITLE OF INVENTION: COMPOSITIONS, TEST KITS AND METHODS FOR DETECTING
; TITLE OF INVENTION: HELICOBACTER PYLORI
; FILE REFERENCE: 626 06 PA
; CURRENT APPLICATION NUMBER: US/10/080,113
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-10-080-113-1

Query Match          13.0%; Score 86; DB 12; Length 249;
Best Local Similarity 29.1%; Pred. No. 0.68;
Matches 34; Conservative 26; Mismatches 39; Indels 18; Gaps 8;

QY 26 YQNEFD---FLMERIHEQIKKGELALFYLOEQINHEFEKPKTKEMKDK--IVAEMDTIIA 80
DB 77 YQKEFEKSLFLQSSSLE--RKG-----YVSQPKDASEIP-QDIKEKALLVLRMDGNTA 128
QY 81 MIDGVRGVLDRMQKDLDFEQY-NLEMLK-KSGDILER---DLKKEARVKKIEV 132
DB 129 ILEDIVESDALSEEKVIDMSSGYLNLNFVEPKSEDIHHSFGIDVSKIKAVIERVEL 185

RESULT 5
US-10-100-699-2
; Sequence 2, Application US/10100699
; Publication No. US20030157650A1
; GENERAL INFORMATION:
; APPLICANT: Lanar, David E.
; APPLICANT: Dutta, Sheetalj
; APPLICANT: Ware, Lisa A.
; TITLE OF INVENTION: Recombinant P. vivax Merozoite Protein-1 p42 Vaccine
; FILE REFERENCE: 003/248/SAP
; CURRENT APPLICATION NUMBER: US/10/100,699
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: US 60/277,002
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 2
; LENGTH: 379
; TYPE: PRT
; ORGANISM: P. vivax
US-10-100-699-2

Query Match          12.4%; Score 82; DB 12; Length 379;
Best Local Similarity 25.0%; Pred. No. 2.9;
Matches 40; Conservative 26; Mismatches 56; Indels 38; Gaps 6;

QY 8 FVATLAVMTVSGEDKKHDYQNEFDLIMERIHEQIKKGELALFYLOEQINHEFE----- 60
DB 28 YLKPLAGMVKTKKQLEHNHNAFNITDMLDSRLKRNRYFLVLSNLPKPKYSSSGEY 87
QY 61 --EPTK---EMKDKIVAEMDTIAMI-----DQVR-----GVLDRL 92
DB 88 IKQFYKLLDLKKKKKLSGYKIGASIDMDLATANDGVTYNNKMGELYKTHLDGVKTEI 147
QY 93 MQRKDLDFEQYNLEMLKSGDILERDLKKEARVKKIEV 132
DB 148 KQVED-DIKQD--EELKLGNVNSQDSKKNEFTAKKAE 184
```

```

RESULT 6
US-09-134-333-13
; Sequence 13, Application US/09134333
; Patent No. US20020076403A1
; GENERAL INFORMATION:
; APPLICANT: LONGACRE-ANDRE, SHIRLEY
; APPLICANT: ROTH, CHARLES
; APPLICANT: NATO, FARIDABANO
; APPLICANT: BARNWELL, JOHN
; APPLICANT: MENDIS, KAMINI
; TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF
; FILE REFERENCE: 0660-0135-0XCIP
; CURRENT APPLICATION NUMBER: US/09/134,333
; CURRENT FILING DATE: 1999-04-18
; EARLIER APPLICATION NUMBER: PCT/FR97/00290
; EARLIER FILING DATE: 1997-02-14
; EARLIER APPLICATION NUMBER: FR96/01822
; EARLIER FILING DATE: 1996-02-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Plasmodium vivax-like sp.
; FEATURE:
; OTHER INFORMATION: Amino Acids 1-140-REGION I
; FEATURE:
; OTHER INFORMATION: Amino Acids 141-178-REGION II
; FEATURE:
; OTHER INFORMATION: Amino Acids 179-283-REGION III
; FEATURE:
; OTHER INFORMATION: Amino Acids 284-380-REGION IV
; OTHER INFORMATION: Amino Acids 284-380-REGION IV
US-09-134-333-13

Query Match      12.4%; Score 82; DB 9; Length 380;
Best Local Similarity 25.0%; Pred. No. 2.9;
Matches 40; Conservative 26; Mismatches 56; Indels 38; Gaps 6;

QY 8 FVATLAVMTVSSEDKKHQYQNEPFDLLMERIHQIKKGELALFYLOEQINHFE-----60
DB 28 YLKPLAGMYTKTKQLQENHVNAFNTITDMLSRKKRNYFLEVLNSDLNPFKYSSGEY 87
QY 61 --EKPTK----EMKDKIVAEMDTIIAMI-----DGVY-----GVLDEL 92
DB 88 IIKDPYKLLDLBKKKLLGISYKVIIGASIDMDLATANDGVTYNNKGELYKTHLDGVKTEI 147
QY 93 MORKDLDIPEQYNLEMLKSGDILERDLKKEEARVKKIEV 132
DB 148 KKVED-DIKKQD--BELKGLGNVNSQDSKKNEFIKKAEL 184

RESULT 7
US-09-742-096-3
; Sequence 3, Application US/09742096
; Patent No. US2002015541A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 200773USODIV
; CURRENT APPLICATION NUMBER: US/09/742,096
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/973,642
; PRIOR FILING DATE: 1998-02-06
; PRIOR APPLICATION NUMBER: PCT/FR96/00894
; PRIOR FILING DATE: 1996-06-12
; PRIOR APPLICATION NUMBER: FR 95/07007
; PRIOR FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: P. falciparum
US-09-742-096-3

Query Match      12.4%; Score 82; DB 10; Length 1786;
Best Local Similarity 24.8%; Pred. No. 21;
Matches 27; Conservative 29; Mismatches 41; Indels 12; Gaps 4;

QY 28 NEFDLLMERIHQIKKGELALFYLOEQINHFEKPTKEMKDKIVAEMDTIIAMID---G 84
DB 1237 NDITSKLIETQE-LNEVEADLIKDMKELKEKALSSEDSKEIIDAQDSTLEKVEIEHD 1295
QY 85 VEGVLDLWQKDLDIPEQYNLEMLKSGDI--LEEDLKKEEARVKKIE 131
DB 1296 ITTLDVVELKQDVE-----EDKIEKVSCLKDLEEDILKEVKEIKELE 1338

RESULT 8
US-10-256-250-16
; Sequence 16, Application US/10256250
; Publication No. US20030170755A1
; GENERAL INFORMATION:
; APPLICANT: Schmitt, Michael
; TITLE OF INVENTION: Tumor-Associated Antigen RHAMM
; FILE REFERENCE: L0461/70143 (JRV/MXA)
; CURRENT APPLICATION NUMBER: US/10/256,250
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/324,989
; PRIOR FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 676
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-256-250-16

Query Match      12.2%; Score 80.5; DB 12; Length 676;
Best Local Similarity 23.6%; Pred. No. 8.6;
Matches 35; Conservative 26; Mismatches 34; Indels 53; Gaps 6;

QY 20 EDKHDYON-----EPDFLLMERIHQIKKGELALFYLOEQ-----55
DB 298 ETEKEDHVNREHNENLNAMQNLQKFILEQREHEKIQKQELQDLSLQOQKELSSSL 357
QY 56 ----INHFEKPTKE---MKDKIVAEMDTIIAMIDGVGVLDLMQR-----KOLDIF 101
DB 358 HQKLCSPQEQEMVYKKNLFEELKQTLDE-----LDKLOQKEQEARLVKQLEEE 406
QY 102 EQYNLEMLKSGDILERDLKKEEARVKK 129
DB 407 AKSRAEELK----LLEEKLGKGEALEK 430

RESULT 9
US-10-256-250-15
; Sequence 15, Application US/10256250
; Publication No. US20030170755A1
; GENERAL INFORMATION:
; APPLICANT: Schmitt, Michael
; TITLE OF INVENTION: Tumor-Associated Antigen RHAMM
; FILE REFERENCE: L0461/70143 (JRV/MXA)
; CURRENT APPLICATION NUMBER: US/10/256,250
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/324,989
; PRIOR FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 709
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-256-250-15
```

```
/ ORGANISM: Homo sapiens
US-10-256-250-15

Query Match      12.2%; Score 80.5; DB 12; Length 709;
Best Local Similarity 23.6%; Pred. No. 9.1;
Matches 35; Conservative 26; Mismatches 34; Indels 53; Gaps 6;

QY 20 EDKKHQYON-----EFDFLLMERIHEQIKKGELALFYLOEQ-----55
Db 282 ETEKEDHVNRRHNENLNAMONLEQKFLQREHEKLOQKELQDLSLQKEKELSSSL 341
QY 56 ---INHFEKPTKE---MKDKIVAEMDTIIMIDGVRGVLDRLMQR-----KLDIF 101
Db 342 HOKLCSFQEMVKKNLFEELKQTLDE-----LDKLOQKEEQEARLVKQLEEE 390
QY 102 EQYNLEMLKSGDILERDLKKEEARVK 129
Db 391 AKSRABELK-----LLEKLGKKEALEK 414

RESULT 10
US-09-978-309A-47
; Sequence 47, Application US/09978309A
; Publication No. US20030100490A1
; GENERAL INFORMATION:
; APPLICANT: Cruz, Tony
; APPLICANT: Pasttrak, Aleksandra
; APPLICANT: Turley, Eva A.
; TITLE OF INVENTION: Compositions and Methods for Treating Cellular Response to
; TITLE OF INVENTION: Injury and Other Proliferating Cell Disorders Regulated by
; TITLE OF INVENTION: Hyaladherin and Hyaluronans
; FILE REFERENCE: 033352-010
; CURRENT APPLICATION NUMBER: US/09/978,309A
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 09/685,010
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: US 09/541,522
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 60/127,457
; PRIOR FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 725
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-978-309A-47

Query Match      12.2%; Score 80.5; DB 11; Length 725;
Best Local Similarity 23.6%; Pred. No. 9.4;
Matches 35; Conservative 26; Mismatches 34; Indels 53; Gaps 6;

QY 20 EDKKHQYON-----EFDFLLMERIHEQIKKGELALFYLOEQ-----55
Db 298 ETEKEDHVNRRHNENLNAMONLEQKFLQREHEKLOQKELQDLSLQKEKELSSSL 357
QY 56 ---INHFEKPTKE---MKDKIVAEMDTIIMIDGVRGVLDRLMQR-----KLDIF 101
Db 358 HOKLCSFQEMVKKNLFEELKQTLDE-----LDKLOQKEEQEARLVKQLEEE 406
QY 102 EQYNLEMLKSGDILERDLKKEEARVK 129
Db 407 AKSRABELK-----LLEKLGKKEALEK 430

RESULT 11
US-10-256-250-14
; Sequence 14, Application US/10256250
; Publication No. US20030170755A1
; GENERAL INFORMATION:
; APPLICANT: Schmitt, Michael
; TITLE OF INVENTION: Tumor-Associated Antigen RHAMM
; FILE REFERENCE: L0461/70143 (JRV/MXA)
```

```
/ CURRENT APPLICATION NUMBER: US/10/256,250
/ CURRENT FILING DATE: 2002-09-26
/ PRIOR APPLICATION NUMBER: US 60/324,989
/ PRIOR FILING DATE: 2001-09-26
/ NUMBER OF SEQ ID NOS: 16
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 14
/ LENGTH: 725
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-256-250-14

Query Match      12.2%; Score 80.5; DB 12; Length 725;
Best Local Similarity 23.6%; Pred. No. 9.4;
Matches 35; Conservative 26; Mismatches 34; Indels 53; Gaps 6;

QY 20 EDKKHQYON-----EFDFLLMERIHEQIKKGELALFYLOEQ-----55
Db 298 ETEKEDHVNRRHNENLNAMONLEQKFLQREHEKLOQKELQDLSLQKEKELSSSL 357
QY 56 ---INHFEKPTKE---MKDKIVAEMDTIIMIDGVRGVLDRLMQR-----KLDIF 101
Db 358 HOKLCSFQEMVKKNLFEELKQTLDE-----LDKLOQKEEQEARLVKQLEEE 406
QY 102 EQYNLEMLKSGDILERDLKKEEARVK 129
Db 407 AKSRABELK-----LLEKLGKKEALEK 430

RESULT 12
US-10-082-830-260
; Sequence 260, Application US/10082830
; Publication No. US20030077604A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Herve
; APPLICANT: Salceda, Susana
; APPLICANT: Liu, Chenghua
; APPLICANT: Turner, Leah
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific
; TITLE OF INVENTION: Genes and Proteins
; FILE REFERENCE: DEX-0249
; CURRENT APPLICATION NUMBER: US/10/082,830
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,802
; PRIOR FILING DATE: 2000-10-27
; NUMBER OF SEQ ID NOS: 282
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 260
; LENGTH: 2383
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-082-830-260

Query Match      12.2%; Score 80.5; DB 15; Length 2383;
Best Local Similarity 22.4%; Pred. No. 42;
Matches 28; Conservative 30; Mismatches 40; Indels 27; Gaps 4;

QY 25 DYQNEFDLIMER-----IHEQIKKGELALFYLOEQINHFEE--XPTKEMDK 70
Db 355 DYQADKALTIVRSVLTERRRQAVQDLRQQLAGCQEAVALNQHQHQQHWEEGKALRQLQK 414
QY 71 IVAEMDTIIMIDGVRGVLDRLMQRKDLDFEYNLEMLKSGDILERD---LKKEEARV 127
Db 415 LTGERDTLAGQVTDLQGEVDSLKER-----ELLQKAREELRQQLVLEQEAURL 464
QY 128 KKIEV 132
Db 465 RAVNV 469
```

```
RESULT 13
US-10-171-311-164
```

```
; Sequence 164, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoerish, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 164
; LENGTH: 1938
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-171-311-164

Query Match      12.1%; Score 80; DB 15; Length 1938;
Best Local Similarity 25.6%; Pred. No. 36;
Matches 33; Conservative 29; Mismatches 47; Indels 20; Gaps 7;

QY 9 VATLAVMTVSGEDKKHYQNEFDLLMERIHQIKKGELALFYLOEQINHPFEKPT----64
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1366 ISTLNQLSDSKKQLQDFAS-----TVEALEBEGKKRFQKEIENLTQQ---YEEKAAAYDK 1417
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 65 -KEMDKIVAEMDTIAMDGVGRVLDRL--MQRK-DLDIFEQYNLEMLKKSGLDILERDL 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1425 LEKTKNRLQQELDLVDLDNQRQLVSNLEKKQKRFQDLAEEKNIS--SKYAD--ERDR 1480
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 KKEEARVKK 129
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1481 AEAEAREXE 1489
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
US-10-171-311-162
; Sequence 162, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoerish, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 162
; LENGTH: 1972
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-171-311-162

Query Match      12.1%; Score 80; DB 15; Length 1972;
Best Local Similarity 25.6%; Pred. No. 37;
Matches 33; Conservative 29; Mismatches 47; Indels 20; Gaps 7;

QY 9 VATLAVMTVSGEDKKHYQNEFDLLMERIHQIKKGELALFYLOEQINHPFEKPT----64
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1366 ISTLNQLSDSKKQLQDFAS-----TVEALEBEGKKRFQKEIENLTQQ---YEEKAAAYDK 1417
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 65 -KEMDKIVAEMDTIAMDGVGRVLDRL--MQRK-DLDIFEQYNLEMLKKSGLDILERDL 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1418 LEKTKNRLQQELDLVDLDNQRQLVSNLEKKQKRFQDLAEEKNIS--SKYAD--ERDR 1473
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 KKEEARVKK 129
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1474 AEAEAREXE 1482
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Search completed: November 17, 2003, 08:29:58
Job time : 30 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2003, 08:22:19 ; Search time 22 Seconds
(without alignments)
253.865 Million cell updates/sec

Title: US-09-877-160A-1

Perfect score: 662

Sequence: 1 MKFIIAFFVATLAVMTSGE.....GDILRDLKKEARVKKIEV 132

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*

1: /cgn2_6/prodata/1/aaa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/aaa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/aaa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/aaa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/aaa/PCUTUS_COMB.pep:*
6: /cgn2_6/prodata/1/aaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	82.5	12.5	282	4	US-09-107-532A-3671
2	82.5	12.5	342	4	US-09-107-532A-5664
3	82	12.4	1786	3	US-08-973-462-8
4	81.5	12.3	344	6	5210183-2
5	81.5	12.3	683	6	5210183-3
6	80	12.1	816	2	US-08-533-306A-6
7	80	12.1	816	2	US-08-742-923A-6
8	80	12.1	885	2	US-08-533-306A-4
9	80	12.1	885	2	US-08-742-923A-4
10	79	11.9	306	4	US-09-107-532A-4841
11	79	11.9	319	4	US-09-107-532A-4723
12	79	11.9	413	4	US-09-198-452A-229
13	78	11.8	142	4	US-09-328-352-6620
14	78	11.8	984	1	US-08-242-932-2
15	78	11.8	984	1	US-08-714-481-2
16	78	11.8	984	5	PCT-US95-06111-2
17	77.5	11.7	475	4	US-09-370-838-193
18	76.5	11.6	107	1	US-08-182-175A-105
19	76.5	11.6	107	1	US-08-474-633A-92
20	76.5	11.6	107	4	US-08-823-771-92
21	76.5	11.6	107	5	PCT-US92-06412-105
22	76	11.5	187	4	US-09-328-352-5117
23	76	11.5	431	4	US-09-286-981B-3
24	75	11.3	1010	4	US-09-134-001C-5178
25	74	11.2	281	4	US-09-914-259-63
26	74	11.2	885	4	US-09-107-532A-5104
27	73.5	11.1	251	4	US-09-286-981B-4

28 73.5 11.1 414 4 US-09-286-981B-10 Sequence 10, Appl
29 73.5 11.1 496 3 US-08-740-223A-16 Sequence 16, Appl
30 73.5 11.1 496 4 US-09-709-188-16 Sequence 16, Appl
31 73.5 11.1 564 3 US-09-308-022-6 Sequence 6, Appl
32 73.5 11.1 631 3 US-08-847-065-25 Sequence 25, Appl
33 73.5 11.1 835 2 US-08-968-751-4 Sequence 4, Appl
34 73 11.0 258 4 US-09-198-452A-563 Sequence 563, Appl
35 73 11.0 284 4 US-09-167-206-10 Sequence 10, Appl
36 73 11.0 284 4 US-09-914-259-40 Sequence 40, Appl
37 73 11.0 284 4 US-09-914-259-48 Sequence 48, Appl
38 73 11.0 284 4 US-09-914-259-50 Sequence 50, Appl
39 73 11.0 285 1 US-08-149-809-24 Sequence 24, Appl
40 72 10.9 388 2 US-08-705-868-4 Sequence 4, Appl
41 72 10.9 388 3 US-09-123-615-4 Sequence 4, Appl
42 72 10.9 421 4 US-09-583-173B-52 Sequence 52, Appl
43 72 10.9 1326 4 US-09-252-991A-17932 Sequence 17932, A
44 72 10.9 2482 1 US-08-328-254-6 Sequence 6, Appl
45 72 10.9 3248 1 US-08-353-700-1 Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-09-107-532A-3671

; Sequence 3671, Application US/09107532A

; Patent No. 6583275

; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

; NUMBER OF SEQUENCES: 7310

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENOME THERAPEUTICS CORPORATION

; STREET: 100 Beaver Street

; CITY: Waltham

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02354

; COMPUTER READABLE FORM:

; MEDIUM TYPE: CD-ROM ISO9660

; COMPUTER: PC

; OPERATING SYSTEM: <Unknown>

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/107,532A

; FILING DATE: 30-Jun-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/085,598

; FILING DATE: 14 May 1998

; APPLICATION NUMBER: 60/051571

; FILING DATE: July 2, 1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Ariniello, Pamela Denise

; REGISTRATION NUMBER: 40,489

; REFERENCE/DOCKET NUMBER: GTC-012

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (781)893-5007

; TELEFAX: (781)893-8277

; INFORMATION FOR SEQ ID NO: 3671:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 282 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: YES

; ORIGINAL SOURCE:

; ORGANISM: Enterococcus faecium

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (B) LOCATION 1...282

; SEQUENCE DESCRIPTION: SEQ ID NO: 3671:

US-09-107-532A-3671

APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS


```
; NAME/KEY: SITE
; LOCATION: 1...413
; OTHER INFORMATION: Xaa=unknown or other
US-09-198-452A-229

Query Match      11.9%; Score 79; DB 4; Length 413;
Best Local Similarity 26.9%; Pred. No. 1.4;
Matches 39; Conservative 22; Mismatches 52; Indels 32; Gaps 7;

QY 2 KFIIA--FFVATLAVMTVSGEDKHDYQNEFDLLMERIHQIKKGLALFYLOEQ---- 55
Db 140 KFLIVDRDFELTRNFTTFGEQVKGIOGNIQD-----LHEKSSLYLELYRLKDLQVL 193
QY 56 INHFEKPTEKMDKIVAEMDTIAMIADGVGVLDRLMQRKD-LDIFEQYNLEML----- 109
Db 194 LNFELPPG-----ILKVD--YDEIEAIKGLFIRLTSLRLDKLDVKAQERKKFINEMSR 244
QY 110 -----KXSGDILLERDLKKEEARVK 129
Db 245 EFKEVEXAFDIVDRATKMLMDRAKK 269

RESULT 13
US-09-328-352-6620
; Sequence 6620, Application US/09328352
; Patent No. 6562958
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6620
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6620

Query Match      11.8%; Score 78; DB 4; Length 142;
Best Local Similarity 25.3%; Pred. No. 0.46;
Matches 25; Conservative 20; Mismatches 42; Indels 12; Gaps 3;

QY 33 LLMERIHE-QIKKGLALFYLOEQINHFEEKPTKEMDKIVAEMDTIAMIADGVGVLDRL 91
Db 8 LILKTFHEYFKEGLDGLFVLSGNF-----LKEIYPTIMLEVDQDTAF-----MAKLT 56
QY 92 LMQRKLDLIFEQYNLEMLKXSGDILLERDLKKEEARVKKI 130
Db 57 LBSGDISLFYNLFNFDSSRDGELLTGSABEQIKLKKV 95

RESULT 14
US-08-242-932-2
; Sequence 2, Application US/08242932
; Patent No. 5595740
; GENERAL INFORMATION:
; APPLICANT: Brady, L. Jeannine
; TITLE OF INVENTION: Cloning of No. 5595740-IgA Fc Binding Forms of
; TITLE OF INVENTION: the Group B Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,481
; FILING DATE: 16-MAY-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/242,932
; FILING DATE: 16-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UFI142
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 2:
```

```
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/242,932
; FILING DATE: 16-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UFI142
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 984 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-242-932-2

Query Match      11.8%; Score 78; DB 1; Length 984;
Best Local Similarity 17.5%; Pred. No. 5.6;
Matches 30; Conservative 37; Mismatches 50; Indels 54; Gaps 4;

QY 16 TVSGEDKXHDYQ-----NEFDLLMERIHQIKKGLA----- 48
Db 88 TILSHQKNEFKTKIDETWSDALLELENOFNETNLLHKHVEEKDKKQKQKTLKQ 147
QY 49 -----LFYLOEQINHFEEKPTKEMDKIVA--EMDTIAMIADGVGVLDRLMQRKLDI 100
Db 148 SDTKVDSLNDKELNHQKQVEKVAQKGTINEDKDSMLKIEDIRKQQAQAKKEDAEV 207
QY 101 FEQYNLEMLKXSGDILE-----RDLKKEEARVKKIEV 132
Db 208 KVLSEEAHSLKQVVEDFRKFKTSEQVTPKKRVKRDLAANNQKIEL 258

RESULT 15
US-08-714-481-2
; Sequence 2, Application US/08714481
; Patent No. 5766606
; GENERAL INFORMATION:
; APPLICANT: Brady, L. Jeannine
; TITLE OF INVENTION: Cloning of No. 5766606-IgA Fc Binding Forms of
; TITLE OF INVENTION: the Group B Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,481
; FILING DATE: 16-MAY-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/242,932
; FILING DATE: 16-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UFI142
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 2:
```

; SEQUENCE CHARACTERISTICS:

LENGTH: 984 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-714-481-2

Query Match 11.8%; Score 78; DB 1; Length 984;

Best Local Similarity 17.5%; Pred. No. 5.6;

Mismatches	30;	Conservative	37;	Mismatches	50;	Indels	54;	Gaps	4;
------------	-----	--------------	-----	------------	-----	--------	-----	------	----

QY 16 TVSGEDKKHDYQ-----NEFDLLMERIHEQIKKGELA----- 48

db 88 TILSHEQNEFKZKIDETNDSALLELENOQFNETNRLLIKHQHEVEKDKAKOOKTLKO 147

QY 49 -----LFYLOEQINHFEKPTKEMKDIVA--EMDTIIAMIDGVRGVLDRLMORKDIDT 100

148 SDTKVDSNIDKELNHQKSOVERMAEQKITNEBKDSMLKKIEDIRKQAOADKEDAEV 207

101 FEQYNLEMLKKSGDILE-----RDLKKEEARVKKIEV 132
QY

207 KQVLEEEAHSKLKQVVEFRKKFKTSEQVTPKRVKRDLAANNENOOKIEL 258

Search completed: November 17, 2003, 08:25:34

Job time : 23 secs